

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 8, 2005, 06:49:22 ; Search time 310 Seconds
(without alignments)
467.855 Million cell updates/sec

Title: US-10-092-934-10

Perfect score: 2034

Sequence: 1 MEPSLLPLRLCNGAISAHK.....FIRGGVSPYLSGWSQTPDLR 375

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_16Dec04:*

1: Geneseqp1980s:*

2: Geneseqp1990s:*

3: Geneseqp2000s:*

4: Geneseqp2001s:*

5: Geneseqp2002s:*

6: Geneseqp2003as:*

7: Geneseqp2003bs:*

8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2034	100.0	375	3 AAB01399	Aab01399 Neuron-as
2	2034	100.0	375	5 ABB81538	Abb81538 Neural th
3	2034	100.0	375	5 AAE29142	Aae29142 AD7c-NTP
4	2034	100.0	375	6 ABR63268	Abr63268 AD7c-NTP
5	2034	100.0	375	6 ADA84017	Ada84017 Human POM
6	2034	100.0	375	6 ABO03024	Abu03024 Human neu
7	2034	100.0	375	6 ABB99774	Abb99774 Amino aci
8	2034	100.0	375	6 ABJ19445	Abj19445 AD7C-neur
9	2034	100.0	375	7 ADB37642	Adb37642 Human neu
10	2034	100.0	375	8 ADR14409	Adr14409 Human NF-
11	1415.5	69.6	397	2 AAR95913	Aar95913 Neural th
12	928	45.6	381	4 AAU30235	Aau30235 Novel hum
13	927.5	45.6	382	4 AAU32610	Aau32610 Novel hum
14	927.5	45.6	382	4 AAU31818	Aau31818 Novel hum
15	927.5	45.6	382	4 AAU32707	Aau32707 Novel hum
16	849	41.7	324	4 AAU29573	Aau29573 Novel hum
17	849	41.7	324	4 AAU29579	Aau29579 Novel hum
18	822.5	40.4	411	4 AAG08428	Aag08428 Novel hum
19	739	36.3	317	4 AAU33200	Aau33200 Novel hum
20	735	36.1	290	4 ABB21913	Abb21913 Novel hum
21	734.5	36.1	361	5 ABB68738	Abb68738 Human pro
22	729	35.8	449	5 ABB01057	Abb01057 Ovary cel
23	701.5	34.5	332	5 ADK34410	Adk34410 Novel hum
24	609	29.9	257	4 AAO10294	Aao10294 Human pol
25	590	29.0	276	4 ABB07919	Abb07919 Novel hum

ALIGNMENTS

RESULT 1

AAB01399

ID AAB01399 standard; protein; 375 AA.

XX AAB01399;

AC AAB01399;

XX AAB01399;

DT 20-OCT-2000 (first entry)

XX AAB01399;

DE Neuron-associated protein.

XX AAB01399;

XX AAB01399;

XX AAB01399;

XX AAB01399;

XX AAB01399;

XX AAB01399;

XX AAB01399;

XX AAB01399;

XX AAB01399;

XX AAB01399;

XX AAB01399;

XX AAB01399;

XX AAB01399;

XX AAB01399;

XX AAB01399;

XX AAB01399;

XX AAB01399;

XX AAB01399;

XX AAB01399;

XX AAB01399;

XX AAB01399;

XX AAB01399;

XX AAB01399;

XX AAB01399;

XX AAB01399;

XX AAB01399;

XX AAB01399;

XX AAB01399;

XX AAB01399;

XX AAB01399;

XX AAB01399;

XX AAB01399;

XX AAB01399;

XX AAB01399;

XX AAB01399;

XX AAB01399;

XX AAB01399;

XX AAB01399;

XX AAB01399;

XX AAB01399;

XX AAB01399;

XX AAB01399;

XX AAB01399;

PS Disclosure; Page 143-144; 145pp; English.

XX Human neuron-associated proteins (NEUPAP) can be used for for treating or preventing a disorder associated with decreased expression or activity of NEUPAP. Antagonists of NEUPAP are useful for treating or preventing a disorder associated with increased expression or activity of NEUPAP. NEUPAP or their fragments or derivatives are useful for treating neurological disorder such as epilepsy, ischemic cerebrovascular disease, stroke, cerebral neoplasms, Alzheimer's disease, Pick's disease, Huntington's disease, dementia and Parkinson's disease. NEUPAPs are also useful for treating other demyelinating diseases, bacterial and viral meningitis, prion diseases including kuru, Creutzfeldt-Jakob disease, nutritional and metabolic diseases of the nervous system, neurofibromatosis, other developmental disorders of the central nervous system, cerebral palsy, neuroskeletal disorders, autonomic nervous system disorders, cranial nerve disorders, spinal cord diseases, muscular dystrophy and other neuromuscular disorders, peripheral nervous system disorders, inherited, metabolic, endocrine, and toxic myopathies, mental disorders including mood, anxiety and schizophrenic disorders, a cell proliferative disorder such as actinic keratosis, arteriosclerosis, atherosclerosis, bursitis, cirrhosis, hepatitis, mixed connective tissue disease (MCTD), myelofibrosis, paroxysmal nocturnal haemoglobinuria, cancers of the adrenal gland, bladder, bone, bone marrow, brain, breast, cervix, and an autoimmune/inflammatory disorder such as acquired immunodeficiency syndrome (AIDS), Addison's disease, adult respiratory distress syndrome, allergies, ankylosing spondylitis, amyloidosis, anemia, asthma, Werner syndrome, complications of cancer, hemodialysis, and extracorporeal circulation, viral, bacterial, fungal parasitic, protozoal, and helminthic infections, and trauma. This protein was designated g3002527

XX SQ Sequence 375 AA;

Query Match 100.0%; Score 2034; DB 3; Length 375;
 Best Local Similarity 100.0%; Pred. No. 1e-198;
 Matches 375; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEFSLLLPRLECNCAISAHNRNLRPGSSDSPASASPVAGITGMCTHARLILYFLVMEF 60
 DB 1 MEFSLLLPRLECNCAISAHNRNLRPGSSDSPASASPVAGITGMCTHARLILYFLVMEF 60

QY 61 LHVGQAGLELPTSDPSVSASQARYRTGHARLCLANFCGRNRVSLMCPSPPELKQST 120
 DB 61 LHVGQAGLELPTSDPSVSASQARYRTGHARLCLANFCGRNRVSLMCPSPPELKQST 120

QY 121 CLSLPKCDWYRRAAVPGLFILFFLRHRCPTLTODEVQWCDHSSLOPSTPEIKHPPASQ 180
 DB 121 CLSLPKCDWYRRAAVPGLFILFFLRHRCPTLTODEVQWCDHSSLOPSTPEIKHPPASQ 180

QY 181 VAGTKDMHYTWLIFIFINFLRQSLNSVTQAGVQWRNLGSLQPLPPGFKLFSCPSLLSS 240
 DB 181 VAGTKDMHYTWLIFIFINFLRQSLNSVTQAGVQWRNLGSLQPLPPGFKLFSCPSLLSS 240

QY 241 WDYRRPPRLANFFVFLVEMGTMFARLILISGPCDLPASASQAGITGVSHARLILNFC 300
 DB 241 WDYRRPPRLANFFVFLVEMGTMFARLILISGPCDLPASASQAGITGVSHARLILNFC 300

QY 301 LFEMESHVSTQAGVQWRNLGSLQPLPPGFKLFSCPSLLSSVDYGHLPHPANFCIFIRGG 360
 DB 301 LFEMESHVSTQAGVQWRNLGSLQPLPPGFKLFSCPSLLSSVDYGHLPHPANFCIFIRGG 360

QY 361 VSPYLSGWSQTPDLR 375
 DB 361 VSPYLSGWSQTPDLR 375

RESULT 2
 ABB81538 standard; protein; 375 AA.

XX ABB81538;
 AC
 XX
 DT 02-SEP-2002 (first entry)
 XX

DE Neural thread protein (NTP).

XX Neural thread protein; NTP; Harlil peptide; Alzheimer's disease;
 KW Down's syndrome; neuroectodermal tumour; astrocytoma; glioblastoma;
 XX hypoxia; ischaemia; cerebral infarction.

OS Homo sapiens.

XX WO200234915-A2.

PN 02-MAY-2002.

PD 25-OCT-2001; 2001WO-US042813.

XX 27-OCT-2000; 2000US-00697590.

PR (NYMO-) NYMOX PHARM CORP.

FA Fitzpatrick J, Averback P, Focht MSS, Bibiano R;
 XX N-PSDB; ABN89470.

XX WPI; 2002-507998/54.

XX New Harlil peptide sequences of the Neural Thread Protein, useful in therapeutic assays, e.g. as targets for developing drugs for treating Alzheimer's disease, ischemia or cerebral infarction, or for diagnosing these diseases.

XX Example 1; Fig 1; 53pp; English.

XX The present invention describes a neural thread protein (NTP) peptide having an amino acid sequence selected from ABB81511 to ABB81529 and their homologues, which are referred collectively as Harlil peptides (I). (I) have neuroprotective, nootropic, vasotropic and cerebroprotective activities, and can be used in peptide therapy. The Harlil peptide sequences can be used as analogues for NTP in therapeutic or diagnostic assays by replacing NTP with the peptide in such an assay. The Harlil peptides are also useful as a trap material in a diagnostic or therapeutic assay. Therefore, the Harlil peptides are useful in binding assays, protein and antibody purification, therapeutics or diagnostics. In particular, the peptides are also useful for diagnosing Alzheimer's disease, Down's syndrome, neuroectodermal tumours, astrocytomas, glioblastomas, hypoxia, ischaemia or cerebral infarction. The peptides are also useful as targets for drug development for the treatment of these diseases. The present sequence represents a neural thread protein given in the exemplification of the present invention

XX SQ Sequence 375 AA;

Query Match 100.0%; Score 2034; DB 5; Length 375;
 Best Local Similarity 100.0%; Pred. No. 1e-198;
 Matches 375; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEFSLLLPRLECNCAISAHNRNLRPGSSDSPASASPVAGITGMCTHARLILYFLVMEF 60
 DB 1 MEFSLLLPRLECNCAISAHNRNLRPGSSDSPASASPVAGITGMCTHARLILYFLVMEF 60

QY 61 LHVGQAGLELPTSDPSVSASQARYRTGHARLCLANFCGRNRVSLMCPSPPELKQST 120
 DB 61 LHVGQAGLELPTSDPSVSASQARYRTGHARLCLANFCGRNRVSLMCPSPPELKQST 120

QY 121 CLSLPKCDWYRRAAVPGLFILFFLRHRCPTLTODEVQWCDHSSLOPSTPEIKHPPASQ 180
 DB 121 CLSLPKCDWYRRAAVPGLFILFFLRHRCPTLTODEVQWCDHSSLOPSTPEIKHPPASQ 180

QY 181 VAGTKDMHYTWLIFIFINFLRQSLNSVTQAGVQWRNLGSLQPLPPGFKLFSCPSLLSS 240
 DB 181 VAGTKDMHYTWLIFIFINFLRQSLNSVTQAGVQWRNLGSLQPLPPGFKLFSCPSLLSS 240

QY 241 WDYRRPPRLANFFVFLVEMGTMFARLILISGPCDLPASASQAGITGVSHARLILNFC 300
 DB 241 WDYRRPPRLANFFVFLVEMGTMFARLILISGPCDLPASASQAGITGVSHARLILNFC 300

Qy 301 LFEMESHVSVTQAGVQVQVNLGSLQPLPGLKRFKFCLSLPSNDYGHLPHPHPANFCIFIRGG 360
 Db 301 LFEMESHVSVTQAGVQVQVNLGSLQPLPGLKRFKFCLSLPSNDYGHLPHPHPANFCIFIRGG 360
 Qy 361 VSPYLSGWSQTPDLR 375
 Db 361 VSPYLSGWSQTPDLR 375

RESULT 3

AAE29142

ID AAE29142 standard; protein; 375 AA.

XX

AC AAE29142;

XX

DT 27-JAN-2003 (first entry)

XX

DE AD7c-NTP protein.

XX

KW Neural thread protein; NTP; hyperplasia; hypertrophy; arteriosclerosis;
 KW haemorrhoid; gene therapy; tumour; vascular disease; atherosclerosis;
 KW inflammatory disease; nutritional deficiency disease; genetic disease;
 KW autoimmune disease; metabolic disease; traumatic disease; intoxication;
 KW infectious disease; congenital malformation; enzyme deficiency disease;
 KW amyloid disease; fibrosis disease; storage disease; radiation disease;
 KW poisoning; environmental disease; endocrine disease; protein therapy;
 KW degenerative disease; mechanical disease; AD7c-NTP protein.

XX

OS Unidentified.

XX

PN WO200274323-A2.

XX

PD 26-SEP-2002.

XX

XX 08-MAR-2002; 2002WO-IB001959.

XX

XX 08-MAR-2001; 2001US-0273957P.

XX

PA (AVER/) AVERBACK P.

XX

XX Averbach P;

XX

XX WPI; 2002-759864/82.

XX

XX N-PSDS; AAD46671.

XX

PT Treating a condition in a patient requiring removal or destruction of
 PT cells, such as a benign or malignant tumor of a tissue or an inflammatory
 PT disease, comprises administering a neural thread protein (NTP) or a NTP
 PT gene to a mammal.

XX

XX Example 2; Fig 1; 70pp; English.

XX

CC The invention relates to a method for treating a condition in a patient
 CC requiring removal or destruction of cells. The method involves
 CC administering to a mammal a neural thread protein (NTP), or administering
 CC to a tumour or other target cell a NTP gene, where the expression of the
 CC NTP gene is induced resulting in expression of the NTP protein. The
 CC method and NTP are useful for treating a condition in a patient requiring
 CC removal or destruction of cells, such as a benign or malignant tumour of
 CC a tissue, a hyperplasia, hypertrophy, or overgrowth of a tissue,
 CC preferably tonsillar hypertrophy or prostatic hyperplasia, a virally,
 CC bacterially, or parasitically altered tissue, or a malformation of a
 CC tissue. Other conditions include a cosmetic modification to a tissue,
 CC such as removal of unwanted facial hair, warts or unwanted fatty tissue,
 CC a vascular disease, particularly atherosclerosis or arteriosclerosis,
 CC haemorrhoids, or varicose veins, an inflammatory disease, autoimmune
 CC disease, metabolic disease, hereditary/genetic disease, traumatic disease
 CC or physical injury, nutritional deficiency disease, infectious disease,
 CC congenital malformation, amyloid disease, fibrosis disease, storage
 CC disease, enzyme deficiency disease, poisoning, intoxication, degenerative
 CC disease, radiation disease, environmental disease, endocrine disease or
 CC mechanical disease. The invention is useful in protein therapy and gene
 CC therapy. The present sequence is AD7c-NTP protein

XX SQ Sequence 375 AA;
 Query Match 100.0%; Score 2034; DB 5; Length 375;
 Best Local Similarity 100.0%; Pred. No. 1e-198;
 Matches 375; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 MEFSLLLPRLCNGAISAHRNLRLPGSSDSPASASPVAGITGCTHARLLLYFFLVEMEF 60
 Db 1 MEFSLLLPRLCNGAISAHRNLRLPGSSDSPASASPVAGITGCTHARLLLYFFLVEMEF 60
 Qy 61 LHVQAGLELPTSDDPVSASQARVYTGHHARLCLANFCGRNRVSLMCPSPKQST 120
 Db 61 LHVQAGLELPTSDDPVSASQARVYTGHHARLCLANFCGRNRVSLMCPSPKQST 120
 Qy 121 CLSLPKCWDYRAAVPGLFILFRLHRCPTLTQDEVQWCDHSSLQSTPEIKHPPASASQ 180
 Db 121 CLSLPKCWDYRAAVPGLFILFRLHRCPTLTQDEVQWCDHSSLQSTPEIKHPPASASQ 180
 Qy 181 VAGTKDMHYTWLIFIFINFLRQSLNSVTQAGVQVNRNLGSLQPLPGLKFCPSLLSS 240
 Db 181 VAGTKDMHYTWLIFIFINFLRQSLNSVTQAGVQVNRNLGSLQPLPGLKFCPSLLSS 240
 Qy 241 WDYRRPRLANPFFVLVEMGFTMFARLILISGCDLPASASQAGITGVSHHARLIFNFC 300
 Db 241 WDYRRPRLANPFFVLVEMGFTMFARLILISGCDLPASASQAGITGVSHHARLIFNFC 300
 Qy 301 LFEMESHVSVTQAGVQVQVNLGSLQPLPGLKRFKFCLSLPSNDYGHLPHPANFCIFIRGG 360
 Db 301 LFEMESHVSVTQAGVQVQVNLGSLQPLPGLKRFKFCLSLPSNDYGHLPHPANFCIFIRGG 360
 Qy 361 VSPYLSGWSQTPDLR 375
 Db 361 VSPYLSGWSQTPDLR 375
 RESULT 4
 ABR63268
 ID ABR63268 standard; protein; 375 AA.
 XX
 AC ABR63268;
 XX
 DT 28-AUG-2003 (first entry)
 XX
 DE AD7c-NTP protein.
 DE
 DE Cytostatic; Antibacterial; Immunosuppressive; Antiinflammatory;
 KW neural thread protein; NTP; tumour.
 KW
 OS Unidentified.
 OS
 XX WO2003008443-A2.
 XX
 XX 30-JAN-2003.
 XX
 XX 19-JUL-2002; 2002WO-CA001105.
 XX
 XX 19-JUL-2001; 2001US-0306150P.
 XX
 XX 19-JUL-2001; 2001US-0306161P.
 XX
 XX 18-NOV-2001; 2001US-0331477P.
 XX
 XX (NYMO-) NYMOX CORP.
 XX
 XX Averbach PA;
 XX
 XX WPI; 2003-247999/24.
 XX
 XX Novel neural thread protein peptide, referred as cell death peptide,
 XX useful for treating prostatic hyperplasia, psoriasis, eczema, dermatosis,
 XX atherosclerosis, cosmetic modification to skin, throat, mouth, muscle.
 XX
 XX Disclosure; Fig 1; 77pp; English.
 XX

CC The present invention relates to a neural thread protein (NTP) peptide
 CC referred to as cell death peptide. Thought to be cytostatic,
 CC antibacterial, immunosuppressive and antiinflammatory. It is useful for
 CC treating a condition in a patient requiring removal or destruction of
 CC cells, for treating a condition such as benign or malignant tumor,
 CC inflammatory disease, autoimmune disease and infectious disease. The
 CC peptide useful for treatment is derived from the amino acid sequence for
 CC a pancreatic thread protein. The peptide is conjugated, linked or bound
 CC to a molecule chosen from antibody or its fragment, antibody-like binding
 CC molecule, where the molecule has a higher affinity for binding to a tumor
 CC or other target than binding to other cells. Treatment using NTP peptides
 CC can remove benign tumors with less risk and fewer of the undesirable side
 CC effects of surgery. The present sequence is an NTP amino acid sequence
 XX
 SQ Sequence 375 AA;

Query Match 100.0%; Score 2034; DB 6; Length 375;
 Best Local Similarity 100.0%; Pred. No. 1e-198;
 Matches 375; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MEFSLLPLRLECNCAISAHNRLRLPGSSDSPASAPVAGITGMCTHARLIILYFVLVEMEF 60
 DB 1 MEFSLLPLRLECNCAISAHNRLRLPGSSDSPASAPVAGITGMCTHARLIILYFVLVEMEF 60
 QY 61 LHVGQAGLELPTSDPSVSASQSARVRTGHHARLCLANFCGRNVRVSLMCPSPWSPKQST 120
 DB 61 LHVGQAGLELPTSDPSVSASQSARVRTGHHARLCLANFCGRNVRVSLMCPSPWSPKQST 120
 QY 121 CLSLPKCWYDRAAVPGFLFILFLRHRCPRTLQDEVQWCDHSSLOPSTPEIKHPPASASQ 180
 DB 121 CLSLPKCWYDRAAVPGFLFILFLRHRCPRTLQDEVQWCDHSSLOPSTPEIKHPPASASQ 180
 QY 181 VAGTKDMHYTWLIFIFLNFRLQSLNSVTQAGVQWRNLGSLQPLPGFKLFCPSLLSS 240
 DB 181 VAGTKDMHYTWLIFIFLNFRLQSLNSVTQAGVQWRNLGSLQPLPGFKLFCPSLLSS 240
 QY 241 WDYRPPRLANFFVFLVEMGFTMFARLILISGPCDLPASASQSAGITGVSHHARLIINFNC 300
 DB 241 WDYRPPRLANFFVFLVEMGFTMFARLILISGPCDLPASASQSAGITGVSHHARLIINFNC 300
 QY 301 LFEMESHVSTQAGVQWRNLGSLQPLPGFKLFCPSLLSS 360
 DB 301 LFEMESHVSTQAGVQWRNLGSLQPLPGFKLFCPSLLSS 360
 QY 361 VSPYLSGWSQTPDLR 375
 DB 361 VSPYLSGWSQTPDLR 375

RESULT 5

ID ADA84017

XX ADA84017 standard; protein; 375 AA.

AC ADA84017;

XX 20-NOV-2003 (first entry)

DT Human POM80 protein.

DE human; marker; expressed sequence tag; EST; arabidopsis; tumour;

XX stress-induced phenotype; hyperosmotic stress; colon cancer; immunogen;

KW vaccine.

XX Homo sapiens.

OS WO2002103028-A2.

XX 27-DEC-2002.

PD 30-MAY-2002; 2002WO-1B004189.

XX 30-MAY-2001; 2001US-0293999P.

PR 22-OCT-2001; 2001US-0330457P.

PR 19-FEB-2002; 2002US-0357144P.

XX (BIOM-) BIOMEDICAL CENT.

XX Baranova AV, Yankovsky NK, Kozlov AP, Lobashev AV, Krukovskaya LL;

XX WPI; 2003-175241/17.

XX N-PSDB; ADA84016.

XX Determining if a nucleic acid is a marker for a phenotype/cell type of

PT interest, by global comparison of expressed sequence tags known to be

PT expressed in the phenotype/cell type with all ESTs expressed in normal

XX tissue.

XX Claim 29; Page 398-400; 516pp; English.

XX The invention relates to a novel method for determining if a nucleic acid

CC is a marker for a predetermined phenotype/cell type of interest from a

CC biological species. The method comprises performing a global comparison

CC of a group of expressed sequence tags (ESTs) known to be expressed in the

CC phenotype/cell type of interest with all ESTs expressed in normal tissue

CC in order to identify ESTs that are preferentially expressed in the

CC phenotype/cell of interest. A method of the invention is useful for

CC determining whether a nucleic acid is a marker for a predetermined

CC phenotype or cell type of interest from a biological species, preferably

CC Arabidopsis or human. The cell type of interest is an abnormal cell such

CC as a tumour cell, and the predetermined phenotype is a stress-induced

CC phenotype such as hyperosmotic stress or high salt conditions. A method

CC of the invention is also useful for determining the progression of colon

CC cancer in a human, for detecting a tumour cell, and for regulating or

CC preventing the growth of a tumour cell. An antibody of the invention is

CC useful for detecting the absence or presence of peptides encoded by

CC tumour-associated markers. A polypeptide of the invention is useful as an

CC immunogen for vaccinating an animal. The present sequence represents a

CC tumour-associated antigen of the invention.

XX SQ Sequence 375 AA;

Query Match 100.0%; Score 2034; DB 6; Length 375;
 Best Local Similarity 100.0%; Pred. No. 1e-198;
 Matches 375; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MEFSLLPLRLECNCAISAHNRLRLPGSSDSPASAPVAGITGMCTHARLIILYFVLVEMEF 60
 DB 1 MEFSLLPLRLECNCAISAHNRLRLPGSSDSPASAPVAGITGMCTHARLIILYFVLVEMEF 60
 QY 61 LHVGQAGLELPTSDPSVSASQSARVRTGHHARLCLANFCGRNVRVSLMCPSPWSPKQST 120
 DB 61 LHVGQAGLELPTSDPSVSASQSARVRTGHHARLCLANFCGRNVRVSLMCPSPWSPKQST 120
 QY 121 CLSLPKCWYDRAAVPGFLFILFLRHRCPRTLQDEVQWCDHSSLOPSTPEIKHPPASASQ 180
 DB 121 CLSLPKCWYDRAAVPGFLFILFLRHRCPRTLQDEVQWCDHSSLOPSTPEIKHPPASASQ 180
 QY 181 VAGTKDMHYTWLIFIFLNFRLQSLNSVTQAGVQWRNLGSLQPLPGFKLFCPSLLSS 240
 DB 181 VAGTKDMHYTWLIFIFLNFRLQSLNSVTQAGVQWRNLGSLQPLPGFKLFCPSLLSS 240
 QY 241 WDYRPPRLANFFVFLVEMGFTMFARLILISGPCDLPASASQSAGITGVSHHARLIINFNC 300
 DB 241 WDYRPPRLANFFVFLVEMGFTMFARLILISGPCDLPASASQSAGITGVSHHARLIINFNC 300
 QY 301 LFEMESHVSTQAGVQWRNLGSLQPLPGFKLFCPSLLSS 360
 DB 301 LFEMESHVSTQAGVQWRNLGSLQPLPGFKLFCPSLLSS 360
 QY 361 VSPYLSGWSQTPDLR 375
 DB 361 VSPYLSGWSQTPDLR 375

RESULT 6

ABU03024

ABU03024 standard; protein; 375 AA.

ABU03024;

20-JAN-2003 (first entry)

Human neural thread protein AD7C-NTP.

Neural thread protein; NTP-peptide; AD7C-NTP; surgical excision; transplantation; grafting; chemotherapy; immunotherapy; vaccination; ablation; cryotherapy; laser therapy; phototherapy; gene therapy; radiation; tumor; hyperplasia; hypertrophy; overgrowth of tissue; malformation of tissue; tonsillary hypertrophy; prostatic hyperplasia; cosmetic modification; vascular disease; atherosclerosis; arteriosclerosis; haemorrhoid; varicose vein; inflammatory disease; autoimmune disease; metabolic disease; traumatic disease; physical injury; nutritional deficiency disease; infectious disease; amyloid disease; fibrosis disease; storage disease; poisoning; congenital malformation; enzyme deficiency disease; poisoning; intoxication; environmental disease; radiation disease; endocrine disease; degenerative disease; mechanical disease.

Homo sapiens.

WO200297030-A2.

05-DEC-2002.

24-MAY-2002; 2002WO-CA000759.

25-MAY-2001; 2001US-0293156P.

(NYMO-) NYMOX CORP.

Averback PA;

WPI; 2003-041406/03.

Novel peptides similar in amino acid sequence to neural thread proteins (NTP), useful for treating unwanted cellular proliferations such as malignant tumors and prostatic hyperplasia.

Disclosure; Fig 1; 78pp; English.

The invention describes an NTP-peptide (I) comprising at least one amino acid sequence corresponding to part of the amino acid sequence of a neural thread protein, AD7C-NTP. The invention provides a method of treating a condition requiring removal or destruction of cells of a mammal comprising administering to a mammal, a therapeutic amount of (I). The treatment is administered to the mammal before, during or after surgical excision, transplantation, grafting, chemotherapy, immunotherapy, vaccination, thermal or electrical ablation, cryotherapy, laser therapy, phototherapy, gene therapy and/or radiation. The method is useful for treatment of benign or malignant tumour; hyperplasia, hypertrophy or overgrowth of tissue; virally, bacterially or parasitically altered tissue; malformation of tissue selected from lung, breast, stomach, pancreas, prostate, bladder, bone, ovary, skin, kidney, sinus, colon, intestines, rectum, esophagus, heart, spleen, salivary gland, blood, brain and its coverings, spinal cord, muscle, connective tissue, adrenal, parathyroid, thyroid, uterus, testis, pituitary, reproductive organs, liver, hair, gall bladder, eye, ear, nose, throat, tonsils, mouth and lymph nodes and lymphoid system; tonsillary hyperplasia; prostatic hyperplasia; cosmetic modification to a tissue; vascular disease (atherosclerosis or arteriosclerosis); haemorrhoids; varicose veins; inflammatory disease; autoimmune disease; metabolic disease; hereditary/genetic disease; traumatic disease; physical injury; nutritional deficiency disease; infectious disease; amyloid disease; fibrosis disease; storage disease; congenital malformation; enzyme deficiency disease; poisoning; intoxication; environmental disease; radiation disease; endocrine disease; degenerative disease and mechanical disease. This is the amino acid sequence of the human neural thread protein AD7C-NTP

SQ Sequence 375 AA;

Query Match 100.0%; Score 2034; DB 6; Length 375;

Best Local Similarity 100.0%; Pred. No. 1e-198;

Matches 375; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEFSLLLPRLCEGNGAISAHRLNLRPGSSSDSPASASPVAGITGCTHARLILYFPLVEMEF 60

DB 1 MEFSLLLPRLCEGNGAISAHRLNLRPGSSSDSPASASPVAGITGCTHARLILYFPLVEMEF 60

QY 61 LHVGQAGLELPTSDDESVASQSAARYTGHARLCLANFCGRNVRVSLMCPWSWPELKQST 120

DB 61 LHVGQAGLELPTSDDESVASQSAARYTGHARLCLANFCGRNVRVSLMCPWSWPELKQST 120

QY 121 CLSLPKCWDYRRRAVPGLFLLFELRHRCPTLTQDEVQWCDHSSLQPSPTPEIKHPPASASQ 180

DB 121 CLSLPKCWDYRRRAVPGLFLLFELRHRCPTLTQDEVQWCDHSSLQPSPTPEIKHPPASASQ 180

QY 181 VAGTKDWHYTWLFIFFIFNFRQSLNSVTQAGVQWRNLGSLQPLPGFKLFSCPSLLSS 240

DB 181 VAGTKDWHYTWLFIFFIFNFRQSLNSVTQAGVQWRNLGSLQPLPGFKLFSCPSLLSS 240

QY 241 WDYRRPRLANFFVFLVEMGFTWPFARLILISGCDLPASASQAGITGVSHHARLIENFC 300

DB 241 WDYRRPRLANFFVFLVEMGFTWPFARLILISGCDLPASASQAGITGVSHHARLIENFC 300

QY 301 LFEMESHVTVQAGVQWPNLGSLOPLPGLKRFSCLSLPPSSWDYGHLPHPANFCIFIRGG 360

DB 301 LFEMESHVTVQAGVQWPNLGSLOPLPGLKRFSCLSLPPSSWDYGHLPHPANFCIFIRGG 360

QY 361 VSPYLSGWSQTPDLR 375

DB 361 VSPYLSGWSQTPDLR 375

RESULT 7

ABB99774

ID ABB99774 standard; protein; 375 AA.

XX AC ABB99774;

XX DT 24-MAR-2003 (first entry)

XX DE Amino acid sequence of human neuronal thread protein AD7C-NTP.

XX KW Human; neuronal thread protein; AD7C-NTP; Alzheimer's disease; histone; neurodegeneration; in vivo gene expression; amphipathic compound; gene therapy.

XX OS Homo sapiens.

XX PN WO200299036-A2.

XX PD 12-DEC-2002.

XX PF 28-MAY-2002; 2002WO-US016429.

XX PR 01-JUN-2001; 2001US-00872968.

XX PA (RHOD-) RHODE ISLAND HOSPITAL.

XX PI Wands JR, De La Monte SM;

XX DR WPI; 2003-140605/13.

XX DR N-PSDB; ABZ23236.

XX PT Inducing prolonged in vivo gene expression in mammal by contacting neuronal tissue with composition comprising Alzheimer's disease-associated neural thread protein 7c antisense nucleic acid, histone, amphipathic compound.

XX PS Disclosure; Page 35; 69pp; English.

XX

CC The present sequence represents a human neuronal thread protein AD7c-NTP.
 CC AD7c-NTP is overexpressed in brains with Alzheimer's disease at early and
 CC intermediate stages of neurodegeneration. The expression of AD7c-NTP may
 CC be reduced using the method of the invention. The specification describes
 CC a method for inducing prolonged in vivo gene expression in a mammal. The
 CC method comprises contacting a non-muscular tissue with a composition
 CC comprising a nucleic acid, histone and an amphipathic compound. The
 CC method is useful for inducing prolonged in vivo gene expression in non-
 CC muscular tissue of a mammal, e.g. neuronal tissue, central nervous system
 CC (CNS) tissue, tissue comprising a post-mitotic neuronal cell, cortical
 CC neuronal cell or hippocampal neuronal cell, glial cell, or vascular
 CC endothelial cell. The method is useful in gene therapy applications to
 CC treat Alzheimer's disease, where the composition comprises antisense
 CC AD7c-NTP nucleic acid

XX Sequence 375 AA;

Query Match 100.0%; Score 2034; DB 6; Length 375;
 Best Local Similarity 100.0%; Pred. No. 1e-198;
 Matches 375; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEFSLLPRLECNCAISAHNRNLRPGSSDSPASASPVAGITGMCTHARLILYFFLVEMEF 60
 DB 1 MEFSLLPRLECNCAISAHNRNLRPGSSDSPASASPVAGITGMCTHARLILYFFLVEMEF 60
 QY 61 LHVGQAGLELPTSDPSVSASQSAARYRTGHARLCLANFCGRNRVSLMCPSPWSPKQST 120
 DB 61 LHVGQAGLELPTSDPSVSASQSAARYRTGHARLCLANFCGRNRVSLMCPSPWSPKQST 120
 QY 121 CLSLPKWDYRRAAVPGFLIFLFLHRCTTLTQDEVQWCDHSSLPQSTPEIKHPPASASQ 180
 DB 121 CLSLPKWDYRRAAVPGFLIFLFLHRCTTLTQDEVQWCDHSSLPQSTPEIKHPPASASQ 180
 QY 181 VAGTKDMHYTWLIFIFLFLHRCTTLTQDEVQWCDHSSLPQSTPEIKHPPASASQ 240
 DB 181 VAGTKDMHYTWLIFIFLFLHRCTTLTQDEVQWCDHSSLPQSTPEIKHPPASASQ 240
 QY 241 WDYRPPRLANFFVFLVEMGFTMPARLILISGPCDLPASASQAGITGVSHARLIFNFC 300
 DB 241 WDYRPPRLANFFVFLVEMGFTMPARLILISGPCDLPASASQAGITGVSHARLIFNFC 300
 QY 301 LFEMESHVTVQAGVQWPNLGSLLQPLPPGLKRFSCLSLPSSWDYGHLPHPANFCIFIRGG 360
 DB 301 LFEMESHVTVQAGVQWPNLGSLLQPLPPGLKRFSCLSLPSSWDYGHLPHPANFCIFIRGG 360
 QY 361 VSPYLSGWSQTPDLR 375
 DB 361 VSPYLSGWSQTPDLR 375

RESULT 8
 ABJ19445
 ID ABJ19445 standard; protein; 375 AA.
 AC ABJ19445;

XX 27-MAR-2003 (first entry)

XX AD7C-neural thread protein.

XX Nootropic; neuroprotective; cell death; tissue necrosis; NTP;
 XX neural thread protein; neurodegenerative disorder; Alzheimer's disease.

XX Unidentified.

XX WO200292115-A2.

XX 21-NOV-2002.

XX 16-MAY-2002; 2002WO-CA000712.

XX 16-MAY-2001; 2001US-0290971P.

XX

PA (NYMO-) NYMOX CORP.

XX Averbach PA;

XX WPI; 2003-129234/12.

DR N-PSDB; AAL54224.

XX Preventing and/or inhibiting cell death and/or tissue necrosis in a
 XX tissue for treating a neurodegenerative disorder, e.g. Alzheimer's
 XX disease, by contacting the live tissue with at least one segment of
 XX neural thread proteins (NTP).

PS Disclosure; Fig 1; 60pp; English.

XX The invention relates to a novel method for preventing and/or inhibiting
 XX cell death and/or tissue necrosis in a tissue comprising contacting the
 XX live tissue with at least one segment of neural thread proteins (NTP).
 XX The methods are composition are useful for treating a neurodegenerative
 XX disorder, such as Alzheimer's disease. This sequence represents the AD7C-
 XX NTP protein relating to the invention

XX Sequence 375 AA;

Query Match 100.0%; Score 2034; DB 6; Length 375;
 Best Local Similarity 100.0%; Pred. No. 1e-198;
 Matches 375; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEFSLLPRLECNCAISAHNRNLRPGSSDSPASASPVAGITGMCTHARLILYFFLVEMEF 60
 DB 1 MEFSLLPRLECNCAISAHNRNLRPGSSDSPASASPVAGITGMCTHARLILYFFLVEMEF 60
 QY 61 LHVGQAGLELPTSDPSVSASQSAARYRTGHARLCLANFCGRNRVSLMCPSPWSPKQST 120
 DB 61 LHVGQAGLELPTSDPSVSASQSAARYRTGHARLCLANFCGRNRVSLMCPSPWSPKQST 120
 QY 121 CLSLPKWDYRRAAVPGFLIFLFLHRCTTLTQDEVQWCDHSSLPQSTPEIKHPPASASQ 180
 DB 121 CLSLPKWDYRRAAVPGFLIFLFLHRCTTLTQDEVQWCDHSSLPQSTPEIKHPPASASQ 180
 QY 181 VAGTKDMHYTWLIFIFLFLHRCTTLTQDEVQWCDHSSLPQSTPEIKHPPASASQ 240
 DB 181 VAGTKDMHYTWLIFIFLFLHRCTTLTQDEVQWCDHSSLPQSTPEIKHPPASASQ 240
 QY 241 WDYRPPRLANFFVFLVEMGFTMPARLILISGPCDLPASASQAGITGVSHARLIFNFC 300
 DB 241 WDYRPPRLANFFVFLVEMGFTMPARLILISGPCDLPASASQAGITGVSHARLIFNFC 300
 QY 301 LFEMESHVTVQAGVQWPNLGSLLQPLPPGLKRFSCLSLPSSWDYGHLPHPANFCIFIRGG 360
 DB 301 LFEMESHVTVQAGVQWPNLGSLLQPLPPGLKRFSCLSLPSSWDYGHLPHPANFCIFIRGG 360
 QY 361 VSPYLSGWSQTPDLR 375
 DB 361 VSPYLSGWSQTPDLR 375

RESULT 9

ADB37642
 ID ADB37642 standard; protein; 375 AA.

XX ADB37642;

XX 04-DEC-2003 (first entry)

XX Human neural thread protein AD7c-NTP.

XX Cytostatic; Antitumour; Antipsoriatic; Dermatological;
 XX Antiatherosclerotic; Antiarteriosclerotic; Vasotropic; Antiinflammatory;
 XX Immunosuppressive; Tranquillizer; Antiemetic; Virucide; AD7c-NTP;
 XX neural thread protein; neuritic sprouting.

XX Homo sapiens.

XX

```

PN WO2003008444-A2.
XX
PD 30-JAN-2003.
XX
XX 19-JUL-2002; 2002WO-CA001106.
XX
XX 19-JUL-2001; 2001US-0306150P.
PR 19-JUL-2001; 2001US-0306161P.
PR 16-NOV-2001; 2001US-0331477P.
XX
XX (NYMO-) NYMOX CORP.
XX
XX Averbach PA, Gemmell J;
XX
XX WPI; 2003-248000/24.
XX
XX N-PSDB; ADB37519.
XX
XX Novel Related peptide or AD7c-neural thread peptide, useful for treating
PT unwanted cellular proliferations, glandular hyperplasia, unwanted facial
PT hair, warts and unwanted fatty tissue.
XX
XX Disclosure; Fig 1; 109pp; English.
XX
XX The present invention relates to AD7c-neural thread protein (NTP) and
CC related proteins and peptides (I; ADB37528-ADB37641). The sequences are
CC useful for treating a condition in a patient requiring removal or
CC destruction of cells. The condition can be selected from benign or
CC malignant tumour of a tissue, hyperplasia, hypertrophy or overgrowth of a
CC tissue, virally, bacterially or parasitically altered tissue, or
CC malformation of a tissue, where the tissue is selected from lung, breast,
CC stomach, pancreas, prostate, bladder, bone, ovary, skin, kidney, salivary
CC colon, intestine, stomach, rectum, oesophagus, heart, spleen, salivary
CC gland, blood, brain and its coverings, spinal cord and its coverings.
CC muscle, connective tissue, adrenal, parathyroid, thyroid, uterus, testis,
CC pituitary, reproductive organs, liver, gall bladder, eye, ear, nose,
CC throat, tonsils, mouth, lymph nodes and lymphoid tissue. The condition is
CC preferably tonsillary hypertrophy, prostatic hyperplasia, psoriasis,
CC eczema, dermatosis, cosmetic modification to a tissue (skin, eye, ear,
CC nose, throat, mouth, muscle, connective, hair or breast tissue), vascular
CC disease (atherosclerosis or arteriosclerosis), haemorrhoids, varicose
CC veins, inflammatory disease, autoimmune disease, metabolic disease,
CC hereditary/genetic disease, traumatic disease or physical injury,
CC nutritional deficiency disease, infectious disease, amyloid disease,
CC fibrosis disease, storage disease, congenital malformation, enzyme
CC deficiency disease, poisoning, intoxication, environmental disease,
CC radiation disease, endocrine disease, degenerative disease and mechanical
CC disease. The peptides are useful for treating unwanted cellular
CC proliferations, glandular (e.g. prostate) hyperplasia, unwanted facial
CC hair, warts and unwanted fatty tissue, or for preparing antibodies that
CC recognize and/or bind to Related proteins, Related peptides or NTP
CC peptides. The present sequence is AD7c-NTP, used to illustrate the
CC invention. AD7c-NTP is a -41kd membrane associated phosphoprotein with
CC functions associated with neuritic sprouting.
XX
XX Sequence 375 AA;
SQ
Query Match 100.0%; Score 2034; DB 7; Length 375;
Best Local Similarity 100.0%; Pred. No. 1e-198;
Matches 375; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MEFSLLPRLCNCAISAHNRNLRLPGSSDSPASASPVAGITGMCTHARLLIYFPLVMEF 60
DB 1 MEFSLLPRLCNCAISAHNRNLRLPGSSDSPASASPVAGITGMCTHARLLIYFPLVMEF 60
QY 61 LHVGQAGLELPTSDPVSASQSAARYTGHARLCLANFCGRNVRVSLMCPSPWSPQLQST 120
DB 61 LHVGQAGLELPTSDPVSASQSAARYTGHARLCLANFCGRNVRVSLMCPSPWSPQLQST 120
QY 121 CLSLPKCWDYRRAAVPGLFILFFLRHCPTLTQDEVQWCDHSSLPQSTPEIKHPPASASQ 180
DB 121 CLSLPKCWDYRRAAVPGLFILFFLRHCPTLTQDEVQWCDHSSLPQSTPEIKHPPASASQ 180
QY 181 VAGTKDMHYTWLIFIFNFRQSLNSVTQAGVQWRNLGSLQPLPGFKLFSCPSLLSS 240

```


CC immunomodulator, cerebroprotective, vasotropic, immunosuppressive or
 CC vulnary activity or for gene therapy. The proteins and nucleotides are
 CC useful for diagnosing, preventing, treating, or ameliorating conditions
 CC or diseases associated with the NF-kappaB pathway. The condition is an
 CC immune disorder, an inflammatory disorder, an inflammatory disorder
 CC related to aberrant NF-kappaB regulation, cancer, aberrant apoptosis,
 CC hepatic disorders, Hodgkin's lymphomas, haematopoietic tumours, hyper-IgM
 CC syndromes, hypohidrotic ectodermal dysplasia, X-linked adhidrotic
 CC ectodermal dysplasia, immunodeficiency, viral infections, HIV-1, HTLV-1,
 CC hepatitis B, hepatitis C, EBV, influenza, viral replication, host cell
 CC survival, evasion of immune responses, rheumatoid arthritis, inflammatory
 CC bowel disease, colitis, asthma, atherosclerosis, cachexia, euthyroid sick
 CC syndrome, stroke, EAE, autoimmune disorders, disorders related to hyper
 CC immune activity, disorders related to aberrant acute phase responses,
 CC hypercongenital conditions, birth defects, necrotic lesions, wounds,
 CC organ transplant rejection, conditions related to organ transplant
 CC rejection, disorders related to aberrant signal transduction,
 CC proliferating disorders, cancers and HIV propagation in cells infected
 CC with other viruses. The present sequence is that of a human protein which
 CC is subject to the novel association with the NF-kappaB pathway of the
 CC invention. Note: This sequence does not appear in the specification but
 CC was obtained by the indexer from Genbank.

XX Sequence 375 AA;

Query Match 100.0%; Score 2034; DB 8; Length 375;
 Best Local Similarity 100.0%; Pred. No. 1e-198;
 Matches 375; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEFSLLLPRLCNGAISAHRNLRPGSSDSPASASPVAGITGCMTHARLILYFFLVEMEF 60
 DB 1 MEFSLLLPRLCNGAISAHRNLRPGSSDSPASASPVAGITGCMTHARLILYFFLVEMEF 60
 QY 61 LHVGQAGLELPTSDPSVSASQARYTGHARCLANFCGRNVRSLMCPSPKQST 120
 DB 61 LHVGQAGLELPTSDPSVSASQARYTGHARCLANFCGRNVRSLMCPSPKQST 120
 QY 121 CLSLPKCDYRRRAAVPGLFILFFLRHRCPTLTQDEVQWCDHSSLPQSTPEIKHPPASASQ 180
 DB 121 CLSLPKCDYRRRAAVPGLFILFFLRHRCPTLTQDEVQWCDHSSLPQSTPEIKHPPASASQ 180
 QY 181 VAGTKDMHHYTWLIFIFIFNLRQSLNSVTQAGVQWRNLGSLQPLPGFKLFSCPSLSS 240
 DB 181 VAGTKDMHHYTWLIFIFIFNLRQSLNSVTQAGVQWRNLGSLQPLPGFKLFSCPSLSS 240
 QY 241 WDYRRPPLANFFVFLVEMGFTMFARLLISGPCDLPASASQAGITGVSHHARLIENFC 300
 DB 241 WDYRRPPLANFFVFLVEMGFTMFARLLISGPCDLPASASQAGITGVSHHARLIENFC 300
 QY 301 LFEMESHVTVQGVQWPNLGSLOPLPGPKLRFSCLSLPSWDYDGLPHPPANFCIFIRGG 360
 DB 301 LFEMESHVTVQGVQWPNLGSLOPLPGPKLRFSCLSLPSWDYDGLPHPPANFCIFIRGG 360
 QY 361 VSPILSGWSQTPDLR 375
 DB 361 VSPILSGWSQTPDLR 375

RESULT 11
 AAR95913
 ID AAR95913 standard; protein; 397 AA.
 XX AAR95913;
 AC AAR95913;
 XX 13-NOV-1996 (first entry)
 DT Neural thread protein.
 DE Neural thread protein.
 XX Neural thread protein; NTP; diagnosis; detection; Alzheimer's disease;
 KW neuroectodermal tumour; malignant astrocytoma; monoclonal antibody;
 KW binding fragment.
 XX Homo sapiens.
 OS

XX WO9615272-A1.
 FN 23-MAY-1996.
 PD 14-NOV-1995; 95WO-US017111.
 XX 14-NOV-1995; 94US-00340426.
 PR 14-NOV-1994; 94US-00340426.
 XX (GEO) GEN HOSPITAL CORP.
 PA De La Monte S, Wands JR;
 FI WPI; 1996-259865/26.
 XX N-PSDB; AAT27738.
 DR Detection of neural thread protein in diagnosis of Alzheimer's disease -
 XX also NTP DNA and protein sequences used in gene and anti-sense therapy.
 PT Claim 22; Page 171-172; 238pp; English.
 PS A method for detecting the presence of neural thread protein (NTP) having
 XX a molecular weight of 8, 14, 17, 21, 26 or 42 kD in a human subject
 CC comprises (a) contacting a sample from a human subject that is suspected
 CC of containing the NTP with at least one molecule capable of binding to
 CC the protein; and (b) detecting any of the molecule bound to the protein.
 CC The binding molecule is selected from an antibody free of natural
 CC impurities, a monoclonal antibody or a binding fragment of either of
 CC these. The method may be used for diagnosing the presence of Alzheimer's
 CC disease, neuroectodermal tumours and a malignant astrocytoma in a human
 XX Sequence 397 AA;

Query Match 69.6%; Score 1415.5; DB 2; Length 397;
 Best Local Similarity 74.6%; Pred. No. 1.7e-135;
 Matches 285; Conservative 10; Mismatches 44; Indels 43; Gaps 7;

QY 1 MEFSLLLPRLCNGAISAHRNLRPGSSDSPASASPVAGITGCMTHARLILYFFLVEMEF 60
 DB 1 MEFSLLLPRLCNGAISAHRNLRPGSSDSPASASPVAGITGCMTHARLILYFFLVEMEF 60
 QY 61 LHVGQAGLELPTSDPSVSASQARYTGHARCLANFCGRNVRSLMCP 110
 DB 61 LHVGQAGLELPTSDPSVSASQARYTGHARCLANFCGRNVRSLMCP 110
 QY 111 SWSPKELKOSTCLSLPKCDYRRRAAVPGLFILFFLRHRCPTLTQDEVQWCDHSSLPQSTPE 170
 DB 111 SWSPKELKOSTCLSLPKCDYRRRAAVPGLFILFFLRHRCPTLTQDEVQWCDHSSLPQSTPE 170
 QY 171 ---IKHPPASASQVAGTKDMHHYTWLIFIFNLRQSLNSVTQAGVQWRNLGSLQPLP 226
 DB 171 SSILLPQPP---KVAGTKDMHHYTWLIFIFNLRQSLNSVTQAGVQWRNLGSLQPLP 226
 QY 227 PGFKLFSCPSLSSWDYRRPPLANFFVFLVEMGF---TWFPAR---LILISGCDLPA 278
 DB 227 PGFKLFSCPSLSSWDYRRPPLANFFVFLVEMGFHVRQVDARSLDLVICLPRP----- 281
 QY 279 SASQAGITGVSHHARLIENFCLEFEMESHVTVQGVQWPNLGSLOPLPGPKLRFSCLSLP 338
 DB 282 --PKVGLQDVTPPTARPIFNFCLEFEMESHVTVQGVQWPNLGSLOPLPGPKLRFSCLSLP 339
 QY 339 SSWDYDGLPHPPANFCIFIRGG 360
 DB 340 SSWDYDGLHHYTLIFVFSLEAG 361

RESULT 12
 AAU30235
 ID AAU30235 standard; protein; 381 AA.
 XX AAU30235;
 AC AAU30235;
 XX 18-DEC-2001 (first entry)
 DT

XX		Novel human secreted protein #726.
DE		
XX		
KW		Human; vaccination; gene therapy; nutritional supplement;
KW		stem cell proliferation; haematopoiesis; nerve tissue regeneration;
KW		immune suppression; immune stimulation; anti-inflammatory; leukaemia.
XX		
OS		Homo sapiens.
XX		
PN		WO200179449-A2.
XX		
PD		25-OCT-2001.
XX		
PB		16-APR-2001; 2001WO-US008656.
XX		
PR		18-APR-2000; 2000US-00552929.
PR		26-JAN-2001; 2001US-00770160.
XX		
PA		(HYSE-) HYSEQ INC.
XX		
PI		Tang YT, Liu C, Drmanac RT;
XX		
DR		WPI; 2001-611725/70.
XX		
PT		Nucleic acids encoding a range of human polypeptides, useful in genetic
PT		vaccination, testing and therapy.
XX		
PS		Claim 20; Page 264-265; 765pp; English.
XX		
CC		The invention relates to novel human secreted polypeptides. The
CC		polypeptides and antibodies to the polypeptides are useful for
CC		determining the presence of or predisposition to a disease associated
CC		with altered levels of polypeptide. The polypeptides are also useful for
CC		identifying agents (agonists and antagonists) that bind to them. Cells
CC		expressing the proteins are useful for identifying a therapeutic agent
CC		for use in treatment of a pathology related to aberrant expression or
CC		physiological interactions of the polypeptide. Vectors comprising the
CC		nucleic acids encoding the polypeptides and cells genetically engineered
CC		to express them are also useful for producing the proteins. The proteins
CC		are useful in genetic vaccination, testing and therapy, and can be used
CC		as nutritional supplements. They may be used to increase stem cell
CC		proliferation; to regulate haematopoiesis; and in bone, cartilage, tendon
CC		and/or nerve tissue growth or regeneration; immune suppression and/or
CC		stimulation; as anti-inflammatory agents; and in treatment of leukaemias.
CC		AU29510-AU33304 represent the amino acid sequences of novel human
CC		secreted proteins of the invention
XX		
SQ		Sequence 381 AA;
		Query Match 45.6%; Score 928; DB 4; Length 381;
		Best Local Similarity 55.9%; Pred. No. 1e-85;
		Matches 223; Conservative 27; Mismatches 101; Indels 48; Gaps 14
Qy	4	SLLLPRLCEGNGAISAHNRNLRLPFGSSDSPASASPAGITGMCTHARLILYFFLVEMEFVHV 63
Dd	4	SLLLPRLCEGNGAILAHCNLCISGSSDSPASASQVTGITKCHHTQLI-FVFLVEMGFPHI 62
Qy	64	GQAGLEIPTSDDPSVASQSARYTG--HHARLCIANCGRRNVSLMCPSW----- 112
Dd	63	AQAGLELLTSDPTL-ASQSAGI-TGVNHHAWLF--FCSRDTVSLCYPGMSRVAXSRIT 118
Qy	113	--SPELKQSTCLSLPKCWDYRAAV-PGLFILFLRHRCPTLTQDEVQCDHSSLQP--- 166
Dd	119	ATSPGLKKXFAFSLPSSRDYRHVPHPGNFCIF-----GRDEVSEPC-----WPQWF 164
Qy	167	STPBKIHPPPASASQVAGTKDMHHYTW----LIPIFINFLRQSLNSVTQAQVQWRNLGSL 222
Dd	165	XTPDLRYPPASASQSAEIIIGVSHHTWPQEVFLFNLYLRWSLDVSVAQARVQRDLSL 224
Qy	223	QPLDPGFKLFCPSLLSNDYTR--PPRIANTFVLVEMGFWTWARLILISGPCDLPASAS 281
Dd	225	QAPPRPKFPFCLSPSSWDYRRPPHPANFVFLVETGFTVLARRVLIXPRDLPASAS 284

QY	282	QSAGITGVSHARLIFNFCLEPMESHVS	VTQAGVQWPNLGSLOPLPGLKRFCSGLSPSSW	343
DB	285	QSAGITGVSHHTRLIFNF--FETGTHSV	TAAVQWYTGISLQRTPELKKSSHLLITSNW	342
QY	342	DYGHLPHPAN-----FCIFIRGGVSP	YLSGWSQTPDLR	375
DB	343	DYRCTPPCPNNLFIYLFYFHRDEGSL	CCPGWSXTPELK	381
RESULT 13				
AAU32610	AAU32610 standard; protein; 382 AA.			
XX	AC	AAU32610;		
XX	DT	19-DEC-2001 (first entry)		
XX	XX	Novel human secreted protein #3101.		
DE	DE			
XX	XX	Human; vaccination; gene therapy; nutritional supplement;		
KW	KW	stem cell proliferation; haematopoiesis; nerve tissue regeneration;		
KW	KW	immune suppression; immune stimulation; anti-inflammatory; leukaemia.		
XX	OS	Homo sapiens.		
OS	OS			
PN	PN	W0200179449-A2.		
XX	XX	25-OCT-2001.		
XX	XX	16-APR-2001; 2001WO-US008656.		
PR	PR	18-APR-2000; 2000US-00552929.		
PR	PR	26-JAN-2001; 2001US-00770160.		
XX	PA	(HYSE-) HYSEQ INC.		
XX	PI	Tang YT, Liu C, Drmanac RT;		
XX	PI	WPI; 2001-611725/70.		
DR	DR			
XX	XX	Nucleic acids encoding a range of human polypeptides, useful in genetic		
PT	PT	vaccination, testing and therapy.		
XX	XX			
PS	PS	Claim 20; Page 637; 765pp; English.		
XX	XX	The invention relates to novel human secreted polypeptides. The		
CC	CC	polypeptides and antibodies to the polypeptides are useful for		
CC	CC	determining the presence of or predisposition to a disease associated		
CC	CC	with altered levels of polypeptide. The polypeptides are also useful for		
CC	CC	identifying agents (agonists and antagonists) that bind to them. Cells		
CC	CC	expressing the proteins are useful for identifying a therapeutic agent		
CC	CC	for use in treatment of a pathology related to aberrant expression or		
CC	CC	physiological interactions of the polypeptide. Vectors comprising the		
CC	CC	nucleic acids encoding the polypeptides and cells genetically engineered		
CC	CC	to express them are also useful for producing the proteins. The proteins		
CC	CC	are useful in genetic vaccination, testing and therapy, and can be used		
CC	CC	as nutritional supplements. They may be used to increase stem cell		
CC	CC	proliferation; to regulate haematopoiesis; and in bone, cartilage, tendon		
CC	CC	and/or nerve tissue growth or regeneration; immune suppression and/or		
CC	CC	stimulation; as anti-inflammatory agents; and in treatment of leukaemias		
CC	CC	AAU29510-AAU33304 represent the amino acid sequences of novel human		
CC	CC	secreted proteins of the invention		
XX	XX	Sequence 382 AA;		
SQ	SQ			

[illegible]

	Query Match	45.6%	Score 927.5;	DB 4;	Length 382;
	Best Local Similarity	55.8%;	Pred. No. 1.le-85;	Mismatches 27;	Gaps 14;
	Matches 223;	Conservative	Indels 49;		
Qy	4	SLLPLRECNCAISAHNNLELPGSSDSPASASPAGITGTCIHTARILLYFLVLEMELHV	63	:	:
Dd	4	SLLPLRECNCAIALHCNLTSGSDSPASAOVTGITKCHETQLT-LFVLVEIMGPHHI	62	:	:
Qy	64	GOAGLEPTDDPSVASOSARVYG--HHAIRCLANFCGRNRVSVMCPGSWS-----	113	:	:
Dd	63	AQAGLELLTSDFTL-AQSAGI-TGVNHNAWLF--PCSRDVISLCYCGWSRVAKSRIT	118	:	:
Qy	114	----PELKSTCGLSKPKCDMYRRAAV-PGIILIFFLRHRCTPILTODEVOMCHSSLQP--	166	: :	: :
Dd	119	ATSAPIGGKFACFSLSRRDYRHVPBHGNGCFE-----GREVFSC-----wpgw	164	: :	: :

	QY	--STPEIXHPPPASQAVAGTKUMHHYTW----	LIFIFPNFLRQSLSNSVTQAGVQRNRLGS	221
	D6	:::::	:::	
	D6	165	EXTPLRYPPASASQSAEIIIGVSHHTWPQEVELFNLFYLFRWSLDSVAQAQVRQRDLGS	224

Db	225	LQAPPRFPFCSLPS	SSWDYRPPHPANFVFLVETGFTVARRVLISXPRDLPASA	284
Qy	281	SQSAGITGVSHHARLI	FNFCFLFEMESHVSTQAGVQWPNLGSLOPLPGLKRFCSLPS	340

Qy	341	WDYGHLPHPAN-----FCIFIRGGVSPYLSGWSOTFDLR	375
Dd	343	WDYRCTPPCPPNLFVLFTFYFHRDEGLCCPGXSTPELK	382

RESULT 15
AAU32707
ID AAU32707 standard: protein: 382 AA-

AAU32707;

DT 18-DEC-2001 (first entry)
XX

DE NOVEL HUMAN SECRETED PROTEIN #3198.
XX

KW human; vaccination; gene therapy; nutritional supplement; stem cell proliferation: haematopoiesis: nerve tissue regeneration:

XX
XX

CS homo sapiens.
XX
PN WQ200179449-A2.

25-OCT-2001.

16-APR-2001; 2001WO-US008656.
PF
XX
XX
18-APR-2000; 2000US-00552929.
PR
PR
26-JAN-2001; 2001US-00770160.
XX
XX
XX
(HYSE-) HYSEQ INC.
PA
XX
XX
Tang YT, Liu C, Drmanac RT;
PI
XX
XX
WPI; 2001-611725/70.
XX
XX
XX
Nucleic acids encoding a range of human polypeptides. useful in genetic

PT vaccination, testing and therapy.

XX Claim 20; Page 652; 765pp; English.

XX The invention relates to novel human secreted polypeptides. The polypeptides and antibodies to the polypeptides are useful for determining the presence of or predisposition to a disease associated with altered levels of polypeptide. The polypeptides are also useful for identifying agents (agonists and antagonists) that bind to them. Cells expressing the proteins are useful for identifying a therapeutic agent for use in treatment of a pathology related to aberrant expression or physiological interactions of the polypeptide. Vectors comprising the nucleic acids encoding the polypeptides and cells genetically engineered to express them are also useful for producing the proteins. The proteins are useful in genetic vaccination, testing and therapy, and can be used as nutritional supplements. They may be used to increase stem cell proliferation; to regulate haematopoiesis; and in bone, cartilage, tendon and/or nerve tissue growth or regeneration; immune suppression and/or stimulation; as anti-inflammatory agents; and in treatment of leukaemias. CC AAU29510-AAU3304 represent the amino acid sequences of novel human secreted proteins of the invention

XX SQ Sequence 382 AA;

Query Match		45.6%;	Score 927.5;	DB 4;	Length 382;
Best Local Similarity		55.8%;	Pred. No. 1.1e-85;		
Matches 223;		Conservative 27;	Mismatches 101;	Indels 49;	Gaps 14;
Qy	4	SLLLPRLECGAISAHRNLRIPGSSDSPASPVAGITGCTHARLILYFPLVEMEFHVV 63			
Db	4	SLLLPRLECGAIIAHNCNLGSSDSPASQVTGITGKCHTQLI-FVFLVEMGFHHI 62			
Qy	64	GQAGLELPTSDPSVSASQSARYRTG--HHAHCLANFCGRNRVSLMCPGWS----- 113			
Db	63	AQAGLELLTSDPTL-ASQSAGI-TGVNHAWLFF--FCSRDVTSLCVPGMSRVAXSRIT 118			
Qy	114	-----PELKQSTCLSLPKCDYRRAV-PGLFILFLRHRCPTLTQDEVQWCDHSSLOP-- 166			
Db	119	ATSAPGLKXCFACFLPSRSDYRHPHPGNFCIF-----GRDEVSPC-----WPGW 164			
Qy	167	-STPEIKHPPASASQVAGTKDMHHTW---LIFIFLNFLOSLSVTOAGVQWRNLGS 221			
Db	165	FXTPLRYPASASQSAEIIIGVSHHTWPQEVFLFNLFIYLRWLSDSVAQARVORRDLGS 224			
Qy	222	LQPLPGFKLFCPSLLSSWDYRR-PPRLANFFVFLVEMGFMPARLILISGPCDLPSA 280			
Db	225	LQAPPRFKPFCSLSPSSWDYRPPHPANFFVFLVETGTFLARRVLISXPRDLPSA 284			
Qy	281	SQSAGITGVSHHARLINFCLFEMESHVSVOAGVQWPNLGSLOPLPGLKRPFCISLPSS 340			
Db	285	SQSAGITGVSHHTRLIENF--FETGTHSVTAAVQWYITIGLSLQPTPELKXSSHILTSN 342			
Qy	341	WDYGHLPHPAN-----FCIFIRGGVSPYLSGWSQTPDLR 375			
Db	343	WDYRCTPPCPNPLFIYLFYFHRDEGSLCCPGWSXTPELK 382			

Search completed: September 8, 2005, 10:37:53

Job time : 312 secs

THIS PAGE BLANK (USP 10)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 8, 2005, 08:34:23 ; Search time 71 Seconds
(without alignments)
508.187 Million cell updates

Title: US-10-092-934-10

Perfect score:

Sequence: 1 MEFSLLPRLECNGAISAHR.....FIRGGVSPYLSGWSQTPDLR 375

Scoring table: BLOSUM62

```
scoring cable: BLOSOME2
Gapop 10.0 , Gapext 0.5
```

Searched: 283416 seqs. 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Minimum DB seq	length: 0
Maximum DB seq	length: 2000000000

Post-processing: Minimum Match 0%

Post-processing: Minimum Match 0%
Maximum Match 100%

Maximum Match 100%
Listing first 45 summaries

Database :

PIR 79:★

1: pir1: v

```

1: pir1:
2: pir2:

```

3: pir3:

pir4:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match		Length	DB	ID	Description
		Score	Match				
1	382.5	18.8	673	4	F40201	artifact-warning s	
2	315.5	15.5	613	4	C40201	artifact-warning s	
3	315.5	15.5	627	4	A40201	artifact-warning s	
4	231.5	14.3	557	4	E40201	artifact-warning s	
5	197.5	9.7	579	4	D40201	artifact-warning s	
6	175	8.6	499	2	S65657	alpha-1C-adrenergi	
7	167	8.2	46	2	I54375	gene NF2 protein -	
8	161	7.9	841	1	I78885	serine/threonine-s	
9	157	7.7	39	2	I54374	gene NF2 protein -	
10	151.5	7.4	301	4	B40201	artifact-warning s	
11	139.5	6.9	100	2	A46010	X-linked retinopat	
12	137.5	6.8	79	2	A56194	thromboxane A-2 re	
13	137.5	6.8	407	2	T02670	probable thromboxa	
14	128	6.3	53	2	A24442	integrin beta-1 ch	
15	112	5.5	440	2	A26359	decay-accelerating	
16	106.5	5.2	1125	2	T19193	hypothetical prote	
17	105	5.2	522	2	T08711	gamma-adaptin homo	
18	102.5	5.0	331	2	S59501	interferon recepto	
19	99.5	4.9	331	2	A54295	interferon alpha/b	
20	96.5	4.7	1643	1	RWGNV	genome polypeptid	
21	96	4.7	458	2	B82175	conserved hypothet	
22	92	4.5	964	2	T15342	hypothetical prote	
23	91.5	4.5	542	2	S36508	transcription fact	
24	91.5	4.5	631	1	RWPYF	transcription fact	
25	91.5	4.5	1776	1	RWPFY	genome polypeptid	
26	90.5	4.4	282	1	SAVL64	middle surface ant	
27	90.5	4.4	832	1	S20752	DNA-directed DNA p	
28	90	4.4	431	1	SAVLIC	large surface anti	
29	89	4.4	299	2	F84785	probable xylocluca	

ALIGNMENTS

RESULT 1

KR0001
F40201

artifact-warning sequence (translated ALU class F) - human

C:Species: Homo sapiens (man)

C:\papers\home\papers\manuscript\11-Aug-1995 #text change 19-May-2000

C; Date: 31-Mar-1992
C: Accession: F40201

R:Claverie, J.M.

personal communication, 1992

A:Reference number: A40201

A: Accession: F40201

A;Molecule type: DNA

A;Residues: 1-673 <CLA>

R; Claverie, J.M.

Genomics 12, 838-841, 1992

A;Title: Identifying coding exons by similarity search: Alu-derived and other potentially

A;Reference number: A40200; MUID:92241891; PMID:15726611

A;Contents: annotation

C;Comment: This "warning" entry is a conceptual translation in all 6 reading frames of or

in-frame stop codons are shown as 'X'.

C;Comment: Any significant similarity of a predicted protein sequence to a portion of th

```

Query Match      18.8%; Score 382.5; DB 4; Length 673;
Best Local Similarity 33.2%; Pred. No. 6.4e-26;
Matches 128; Conservative 19; Mismatches 111; Indels 127; Gaps 17;

-
Qy      16 ISAHNRLRPGSSSDSPASASPAGITCMCTHARLLILYFLFVEMEFHVGGAGLELPISDD 75
Db      371 VLAHCSLNLGSSDSPASVSRAVIGTGMGRHSMLI-YVFLTIETQHVDQAGLKLLTSSD 429
          : : : : : | | | | | : : : : : | | | | | : : : : : | | | | |
Qy      76 PSVASQSARYRTGTGHARLCLANFCGRNRVSLMCPMSPE-LKQSTCLSLPKCWDYRRAA 134
Db      430 -----LPSWSPKVLGXQAXATTFSSXXXF---V 453
          : : : : : | | | | | : : : : : | | | | | : : : : : | | | | |
Qy      135 VPLGLFILFFL-----RHRCP---TLTQDEVQMCDHSSLQPSTPEIKHPASA 178
Db      454 FGGFFFFFALFLFXALALTPELECSKFMTAAASTSV-QAILLPLSPVXLQLQAWA 512
          : : : : : | | | | | : : : : : | | | | | : : : : : | | | | |
Qy      179 SQVAGT----KDMHYTWLIPIFIENFLROSLNSVTQAGVQRNLGSLQPLPPGFKLPSG 234
Db      513 A-IPGXPMYFKXSHSFMTLI-----RLVSNXPQV-----ICPPG----- 546
          : : : : : | | | | | : : : : : | | | | | : : : : : | | | | |
Qy      235 PSLLSLDWYER-PRLANPFVELFVEMGFTWFARLLILSGCDLPASASAGITGVSHHA 293
Db      547 --LPKCDDDRREPPHPAXXXLFL---GVFPF----- 572
          : : : : : | | | | | : : : : : | | | | | : : : : : | | | | |
Qy      294 RLIPNFCLFMESHSVTQAGVQW---PNLAGSLOPLPGLKRFSCLSLPSSWDYHGHLPPHP 350
Db      573 --FLLCFCEXDRLPXHG--WSAVASFGLSQPQPPGFKRFSCCLCLPCSWDYRHGPFL 628
          : : : : : | | | | | : : : : : | | | | | : : : : : | | | | |
Qy      351 ANFCIFIRGGVSPYLGSWQSTPDLR 375
          : : : : : | | | | | : : : : : | | | | | : : : : : | | | | |

```

Db 629 ANLCIFNRDVTSPCKXGWSQTDLK 653

RESULT 2

C40201

artifact-warning sequence (translated ALU class C) - human

C:Species: Homo sapiens (man)

C>Date: 31-Mar-1992 #sequence_revision 11-Aug-1995 #text_change 19-May-2000

C:Accession: C40201

R:Claverie, J.M.

personal communication, 1992

A:Reference number: A40201

A:Accession: C40201

A:Molecule type: DNA

A:Residues: 1-613 <CLA>

R:Claverie, J.M.

Genomics 12, 838-841, 1992

A:Title: Identifying coding exons by similarity search: Alu-derived and other potentially

A:Reference number: A40200; MUID:92241891; PMID:1572661

A:Contents: annotation

C:Comment: This "warning" entry is a conceptual translation in all 6 reading frames of c

in-frame stop codons are shown as 'X'.

C:Comment: Any significant similarity of a predicted protein sequence to a portion of th

Query Match 15.5%; Score 315.5; DB 4; Length 613;

Best Local Similarity 33.4%; Pred. No. 5e-20;

Matches 98; Conservative 9; Mismatches 73; Indels 113; Gaps 8;

QY 6 LLPLECNGAISAHNRNLRLPGSSDSPASAPVAGITGCTHARLILYFFLVEMFELHVQ 65

Db 424 LSPRLCNGEISAHCKLHPLGLCHSPAFVCGTGTGARTMPS-XFFVFLVEMGFCHVQ 482

QY 66 AGLELPTSDDPSPVSASQSAARYTGHARLCLANFCGRNRVSLMCPSPSPKQSTCLSLP 125

Db 483 AGLEULAS-----XSTHLCLP 498

QY 126 KCWDYRR-----AAYVGLFILFRLHRCPTLTQDEVQWCHDSSLOPSTPEIKHPASA--- 178

Db 499 KCWDYRRFELHPAXXXFF-----XDRVSLC-----HPGWSAWAR 536

QY 179 SQVAGTKDMHYTWLIFIFNLRQSLNSVTQAGVQWRNLGSLQPLPGPKLFSCPSLL 238

Db 537 SRLTASSTRVYAILL-----POPSAVVGLQAPA-- 565

QY 239 SSWDYRRPRLANFFVFLVEMGFTMFARLILISGCDLPASASQAGITGVSH 291

Db 566 -----PCPANFLYFXKKWGFAMLARLVNSWPHDPPTTASQAGITGVSH 610

RESULT 3

A40201

artifact-warning sequence (translated ALU class A) - human

C:Species: Homo sapiens (man)

C>Date: 31-Mar-1992 #sequence_revision 11-Aug-1995 #text_change 19-May-2000

C:Accession: A40201

R:Claverie, J.M.

personal communication, 1992

A:Reference number: A40201

A:Accession: A40201

A:Molecule type: DNA

A:Residues: 1-627 <CLA>

R:Claverie, J.M.

Genomics 12, 838-841, 1992

A:Title: Identifying coding exons by similarity search: Alu-derived and other potentiall

A:Reference number: A40200; MUID:92241891; PMID:1572661

A:Contents: annotation

C:Comment: This "warning" entry is a conceptual translation in all 6 reading frames of c

in-frame stop codons are shown as 'X'.

C:Comment: Any significant similarity of a predicted protein sequence to a portion of th

Query Match 15.5%; Score 315.5; DB 4; Length 627;

Best Local Similarity 33.4%; Pred. No. 5.2e-20;

Matches 98; Conservative 13; Mismatches 67; Indels 115; Gaps 7;

QY 6 LLPLECNGAISAHNRNLRLPGSSDSPASAPVAGITGCTHARLILYFFLVEMFELHVQ 65

Db 436 LSPRLCNGEISAHCKLHPLGLCHSPAFVCGTGTGARTMPS-XFFVFLVETGFRHSQ 494

QY 66 AGLELPTSDDPSPVSASQSAARYTGHARLCLANFCGRNRVSLMCPSPSPKQSTCLSLP 125

Db 495 DGLDLLT-----SXSAR-----LCLP 510

QY 126 KCWDYRRAAVPGULFILFRLHRCPTLTQDEVQWCHDSSLOPSTPEIKHPASASQVAGTK 185

Db 511 KCWDYRRRETAPG----- 522

QY 186 DMHHYTWLIFIF-----NFLRQSLNSVTQ-----AGVQWRNLGSLQPLPGPKLFSCPS 236

Db 523 ---XXXFLPVFIFRFGVSLCQGWASVARSRLTASSASRVHAILLPQPPKXGLQAPA 579

QY 237 LLSWDYRRPRLANFFVFLVEMGFTMFARLILISGCDLPASASQAGITGV 289

Db 580 L-----RPNANFLYFXRRGFTVVARVMSISXPRDPALASQAGITGV 622

RESULT 4

E40201

artifact-warning sequence (translated ALU class E) - human

C:Species: Homo sapiens (man)

C>Date: 31-Mar-1992 #sequence_revision 11-Aug-1995 #text_change 19-May-2000

C:Accession: E40201

R:Claverie, J.M.

personal communication, 1992

A:Reference number: A40201

A:Accession: E40201

A:Molecule type: DNA

A:Residues: 1-597 <CLA>

R:Claverie, J.M.

Genomics 12, 838-841, 1992

A:Title: Identifying coding exons by similarity search: Alu-derived and other potentiall

A:Reference number: A40200; MUID:92241891; PMID:1572661

A:Contents: annotation

C:Comment: This "warning" entry is a conceptual translation in all 6 reading frames of c

in-frame stop codons are shown as 'X'.

C:Comment: Any significant similarity of a predicted protein sequence to a portion of th

Query Match 14.3%; Score 291.5; DB 4; Length 597;

Best Local Similarity 30.5%; Pred. No. 6.6e-18;

Matches 118; Conservative 24; Mismatches 106; Indels 139; Gaps 20;

QY 5 LLLPRLECNGAISAHNRNLRLPGSSDSPASAPVAGITGCTHARLILYF-----FLVEM 58

Db 310 LLLPRPEYDGTISP--QTLPGPKQSPVSASRVAGITGRHHAQLILYFXRWGFSMLV 367

QY 59 EFLHVQAGLELPTSDDPSPVSASQSAARYT-----GHHARLC 95

Db 368 KLVNSQPVIRPALANSQSAGIT-GMSYHTWXXFFFFFETFCSCCGRSTMAQSHKLC 426

QY 96 LANFCGRNRVSLMCPSPSPKQSTCLSLPKCKDYRRAAVPGLFILFRLHRCPTLTQDE 155

Db 427 ---FLAGSSNLSLQPEX-----LGLQAC-----ATTPSXFIYF-----SRDG 461

QY 156 V-----QWCHDSSLOPSTPEIKHPASASQVAGTKDMHHYTWLIFIFNLRQSLNSVT 210

Db 462 VSPCWSWSQTPNLRXSAPPXT-PKALGLQAXATP-----GXKXFFLRRSFALVA 512

QY 211 QAGVQWRNLGSLQPLPGF-----KLFSCPSLLSSWDYRR--PRLANFFVFLVEMGFTMF 264

Db 513 QAGVRVHNLTA-----NFWASWQAIIISCLSLSPSSWDYRHAPRPAN-FILVEMGFL-- 563

QY 265 ARILISGCDLPASASQAGITGVSHARLIFNFCLEFEMESHVTOAGVQWRNLGSLQ 324

Db 564 -----HVQAGLKLPTSGD--- 577

A;Residues: 424-499 <HIR>
A;Cross-references: EMBL:D32202
C;Genetics:
A;Gene: GDB:ADRALC; ADRAL11
A;Cross-references: GDB:I28088; OMIM:104221
A;Map position: 8p21-8p11.2
C;Superfamily: vertebrate rhodopsin
C;Keywords: alternative splicing; neurotransmitter receptor

Query Match 8.6%; Score 175; DB 2; Length 499;
Best Local Similarity 58.0%; Pred. No. 1.1e-07;
Matches 40; Conservative 5; Mismatches 20; Indels 4; Gaps 1

Qy 9 RLECNGLSAHNRNLRLPGSSDSPASAPVAGITGWC-----THARILLYFFELVEMEFLLHG 64
|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:::|||
Db 431 RLECSGMILAHCNLRLPGRSDPASASOAGTTGDVPGRHRQAQLIFVLVTCTGFHHVG 490
||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:

Qy 65 QAGLELPTS 73
|||:|||:|||
Db 491 QDDLDTLS 499
|||:|||:|||

RESULT 7
I54375
gene NF2 protein - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 01-Nov-1996 #sequence_revision 01-Nov-1996 #text_change 16-Aug-2004
C;Accession: I54375
R:Akakawa, H.; Hayashi, N.; Nagase, H.; Ogawa, M.; Nakamura, Y.
Hum. Mol. Genet. 3, 565-568, 1994
A>Title: Alternative splicing of the NF2 gene and its mutation analysis of bre
A;Reference number: I54375; MUID:94348501; PMID:8069299
A;Accession: I54375
A>Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: mRNA
A;Residues: 1-46 <RES>
A;Cross-references: UNIPROT:Q16230; GB:S73853; NID:G689372; PIDN:AAB31736.1; P
C;Genetics:
A;Gene: GDB:NF2
A;Cross-references: GDB:I20232; OMIM:101000
A;Map position: 22q12.2-22q12.2
C;Superfamily: laminin-type EGF-like homology

Query Match 8.2%; Score 167; DB 2; Length 46;
Best Local Similarity 76.2%; Pred. No. 3.6e-08;
Matches 32; Conservative 1; Mismatches 9; Indels 0; Gaps 0

Qy 302 FEMESHSTQAGVQWPNLGSLQPLPPGLKRFSCSLSPSSWDY 343
| || | || | || | : || | || | || | || | || | || | || |
Db 4 FNCSCSVTLTLAGVQWRDLGLLOLPUPPKFKRFSCSLFPSSWDY 45
| || | || | || | || | || | || | || | || | || | || |

RESULT 8
I78885
serine/threonine-specific protein kinase (EC 2.7.1.-) STK2 - human
C;Species: Homo sapiens (man)
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
C;Accession: I78885
R:Lavedakou, E.N.; He, M.; Baptist, E.W.; Craven, R.J.; Cance, W.G.; Weicsh, P
Oncogene 9, 1977-1988, 1994
A>Title: Two novel human serine/threonine kinases with homologies to the cell
A;Reference number: I58396; MUID:94269838; PMID:8208544
A;Accession: I78885
A>Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: mRNA
A;Residues: 1-841 <RRS>
A;Cross-references: UNIPROT:P51957; GB:L20321; NID:G348244; PIDN:AAA36658.1; P
C;Genetics:
A;Gene: GDB:STK2
A;Cross-references: GDB:374125
A;Map position: 3p21.1-3p21.1
C;Superfamily: human serine/threonine-specific protein kinase STK2; protein ki
C;Keywords: phosphotransferase

F:4-261/Domain: protein kinase homology <KIN>

```

Query Match          7.9%; Score 161; DB 1; Length 841;
Best Local Similarity 46.2%; Pred. No. 3.6e-06;
Matches 42; Conservative 14; Mismatches 25; Indels 10; Gaps 4;

Qy 6 LLPRLECGAISAHRNLRPGSSDSPASASPVAGITGCTHAR-LILYFFILVEMEFHLVG 64
Db 460 LSPKLECSGTILAHNLRLLGSSDSPASASRVAGITGVCHHAQDQVAGECIKQ----G 515

Qy 65 QAGLELP---TSDDPSVAS--QSARYRTGH 90
Db 516 RIHPDLPQHNSEPSLSRQRQRKREOTEH 546

RESULT 9
I54374
gene NF2 protein - human (fragment)
C:Species: Homo sapiens (man)
C>Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 16-Aug-2004
C:Accession: I54374
R:PyKett, M.J.; MurphY, M.; Harnish, P.R.; George, D.L.
Hum. Mol. Genet. 3:559-564, 1994
A>Title: The neurofibromatosis 2 (NF2) tumor suppressor gene encodes multiple alternatively spliced transcripts
A:Reference number: I54374; MUID:94348500; PMID:8069298
A:Accession: I54374
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-39 <RES>
A:Cross-references: UNIPROT:Q16230; GB:L27065; NID:9463120; PIDN:AAA62422.1; PID:9463121
C:Genetics:
A:Gene: GDB:NF2
A:Cross-references: GDB:120232; OMIM:101000
A:Map position: 22q12.2-22q12.2
C:Superfamily: laminin-type EGF-like homology

Query Match          7.7%; Score 157; DB 2; Length 39;
Best Local Similarity 78.9%; Pred. No. 2.3e-07;
Matches 30; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

Qy 306 SHSVTQAGVQWPNLGSLOPLPGLKRFSLPSWDY 343
Db 1 SCSVTLAGVQWRDLGLLOPLPKRFSLPSWDY 38

RESULT 10
B40201
artifact-warning sequence (translated ALU class B) - human
C:Species: Homo sapiens (man)
C>Date: 31-Mar-1992 #sequence_revision 11-Aug-1995 #text_change 19-May-2000
C:Accession: B40201
R:Claverie, J.M.
personal communication, 1992
A:Reference number: A40201
A:Accession: B40201
A:Molecule type: DNA
A:Residues: 1-301 <CLA>
R:Claverie, J.M.
Genomics 12, 838-841, 1992
A>Title: Identifying coding exons by similarity search: Alu-derived and other potential
A:Reference number: A40200; MUID:92241891; PMID:1572661
A:Contents: annotation
C:Comment: This "warning" entry is a conceptual translation in all 6 reading frames of c
in-frame stop codons are shown as 'X',
C:Comment: Any significant similarity of a predicted protein sequence to a portion of th

Query Match          7.4%; Score 151.5; DB 4; Length 301;
Best Local Similarity 22.4%; Pred. No. 7.6e-06;
Matches 66; Conservative 15; Mismatches 45; Indels 169; Gaps 10;

Qy 6 LLPRLECGAISAHRNLRPGSSDSPASASPVAGITGCT-----HARLILYFFILVEMEF 60
Db 1 LSPKLECSGTILAHNLRLLGSSDSPASASRVAGITGVCHHAQDQVAGECIKQ----G 515

```

```

Db 168 MLPKVLNWSQATILLQHP-----KVLGLOMLATMPCXXXTQLILFYF----- 211
Qy 61 LHVGQAGLELPTSDDPVSASQSARYRTGHARCLANFCGRNVSLSMCPWSM--PELKQ 118
Db 212 -----YRAG-----VLLCCPSWSXTPGLKQ 231
Qy 119 STCLSLPKCWDYRAAVPGLFILFFLRHRCPTLTQDEVQWCDHSSLOPSTPEIKHPASA 178
Db 232 SSVSIFPKCDYR-----C-----XPPCPA 251
Qy 179 SQVAGTKDMHHYTLIFIFINFLRQSLNSVTQAGVQWRNLGSLQPLPPGPKLFSCLP 238
Db 252 XXXPSX-----FYFIFELGSCYVAQGLE----- 276
Qy 239 SSNDYRPPRLANFFVLVEMGFTMFARLILISGPCDLPASASOSAGITGVSHHA 293
Db 277 -----LLVSSNP-----PTSASOSAGITDVSHHA 300

RESULT 11
A46010
X-linked retinopathy protein (C-terminal, clone XEH.8c) - human (fragment)
C:Species: Homo sapiens (man)
C>Date: 21-Sep-1993 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004
C:Accession: A46010
R:Wong, P.; MacDonald, I.M.; Sood, R.; Smith, C.; Pilon, R.; Tenniswood, M.
Genomics 15, 467-471, 1993
A>Title: Identification and partial characterization of a candidate gene for X-linked re
A:Reference number: A46010; MUID:93224131; PMID:8468040
A:Accession: A46010
A>Status: preliminary
A:Molecule type: nucleic acid
A:Residues: 1-100 <WON>
A:Cross-references: UNIPROT:Q07826; GB:S58722; NID:9299470; PIDN:AAB26149.1; PID:9299471
A>Note: sequence extracted from NCBI backbone (NCBIN:129339, NCBIPI:129340)

Query Match          6.9%; Score 139.5; DB 2; Length 100;
Best Local Similarity 34.2%; Pred. No. 2.5e-05;
Matches 38; Conservative 6; Mismatches 22; Indels 45; Gaps 2;

Qy 197 FIFNFLRQSLNSVTQAGVQWRNLGSLQPLPFGPKLFSCLPSLLSSWDYRRPRLANFFVL 256
Db 1 FFFFEFETESCSVAEAGVQWCDLGSLSKSPPP----- 31

Qy 257 VEMGFTMFARLILISGPCDLPASASOSAGITGVSHHARLIENFCLFEMESH 307
Db 32 -----GSSDSPASASRVAGITGMHHHTQLIFVF-LVETGSH 66

```

```

RESULT 12
A56194
thromboxane A-2 receptor, endothelial - human (fragment)
C:Species: Homo sapiens (man)
C>Date: 19-Oct-1995 #sequence_revision 19-Oct-1995 #text_change 09-Jul-2004
C:Accession: A56194
R:Raychowdhury, M.K.; Yukawa, M.; Collins, L.J.; McGrail, S.H.; Kent, K.C.; Ware, J.A.
J. Biol. Chem. 270, 7011, 1995
A:Reference number: A56194; MUID:95204505; PMID:7896853
A:Accession: A56194
A>Status: preliminary; not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-79 <RAY>
A:Cross-references: UNIPROT:Q9UCY2; GB:U11271
C:Superfamily: prostaglandin E receptor EPI

Query Match          6.8%; Score 137.5; DB 2; Length 79;
Best Local Similarity 49.3%; Pred. No. 2.8e-05;
Matches 37; Conservative 8; Mismatches 25; Indels 5; Gaps 2;

Qy 6 LLPRLECGAISAHRNLRPGSSDSPASASPVAGITGCTHARLILYFFILVEMEFHLVGQ 65
Db 5 LWPSEYSGTISAHNLRLLFGSSDRASASRAAGITGV-SHCARPCMLFPDFDL----- 59

```

C;Species: Homo sapiens (man)
C;Date: 05-Oct-1988 #sequence_revision 05-Oct-1988 #text_change 09-Jul-2004
C;Accession: A26359; A39702; S16187; S23138; A27258
C;Caras, I.W.; Davitz, M.A.; Rhee, L.; Weddell, G.; Martin Jr., D.W.; Nussenzweig, V.
Nature 325, 545-549, 1987
A;Title: Cloning of decay-accelerating factor suggests novel use of splicing to generate
A;Reference number: A26359; MUID:87115845; PMID:2433596
A;Accession: A26359
A;Molecule type: mRNA
A;Residues: 1-440 <CAR>
A;Cross-references: UNIPROT:P08174; GB:M30142
R;Ewulonu U.K.; Ravi, L.; Medof, M.E.
Proc. Natl. Acad. Sci. U.S.A. 88, 4675-4679, 1991
A;Title: Characterization of the decay-accelerating factor gene promoter region.
A;Reference number: A39702; MUID:91271256; PMID:1711208
A;Accession: A39702
A;Molecule type: DNA
A;Residues: 1-79, 'T', 81-104 <EWU>
A;Cross-references: GB:M64356
A;Note: the authors translated the codon AGT for residue 85 as Met
R;Nakano, Y.; Sugita, Y.; Ishikawa, Y.; Choi, N.H.; Tobe, T.; Tomita, M.
Biochim. Biophys. Acta 1074, 326-330, 1991
A;Title: Isolation of two forms of decay-accelerating factor (DAF) from human urine.
A;Reference number: S16187; MUID:91291869; PMID:1712233
A;Accession: S16187
A;Status: preliminary
A;Molecule type: protein
A;Residues: 35-47 <BIO>
R;Nakano, Y.; Sumida, K.; Kikuta, N.; Miura, N.H.; Tobe, T.; Tomita, M.
Biochim. Biophys. Acta 1116, 235-240, 1992
A;Title: Complete determination of disulfide bonds localized within the short consensus
A;Reference number: S23138; MUID:92305034; PMID:1377029
A;Accession: S23138
A;Status: preliminary
A;Molecule type: protein
R;Sugita, Y.; Negro, T.; Matsuda, T.; Sakamoto, T.; Tomita, M.
J. Biochem. 100, 143-150, 1986
A;Title: Improved method for the isolation and preliminary characterization of human DAF
A;Reference number: A27258; MUID:87008461; PMID:2428813
A;Accession: A27258
A;Molecule type: protein
A;Residues: 35, 'X', 37, 'G', 39-51, 'P', 53-55, 'X', 57-58, 'X', 60-63 <SUG>
A;Note: Gly-35 and Leu-38 were also found
C;Comment: For an alternative splice form, see PIR:B26359
C;Genetics:
A;Gene: GDB:DAF
A;Cross-references: GDB:119088; OMIM:125240
A;Map position: 1q32-1q32
C;Superfamily: decay-accelerating factor; complement factor H repeat homology
C;Keywords: alternative splicing; glycoprotein
F;1-34/Domain: signal sequence #status predicted <SIG>
F;35-440/Product: decay-accelerating factor 1 #status predicted <MAT>
F;36-94/Domain: complement factor H repeat homology <FH01>
F;98-158/Domain: complement factor H repeat homology <FH02>
F;163-220/Domain: complement factor H repeat homology <FH03>
F;225-283/Domain: complement factor H repeat homology <FH04>
F;95/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 5.5%; Score 112; DB 2; Length 440;
Best Local Similarity 20.6%; Pred. No. 0.038;
Matches 87; Conservative 43; Mismatches 100; Indels 192; Gaps 22;

Qy 22 LRLPGSSDSP-ASASPVAGITGCTHARLLLYFFLVEMEFHVQAGLELPTSDPSVS 79
Db 70 VKTPGKDSVILKGSQWSDIEFCNRS-----CEVPTRLN---S 106
Qy 80 ASQSARYRTGHARLCLANFCGRNRSVMC-----PSWSPELKOSTCLSLPKCWDYRR 132
Db 107 ASLKQPIYQNY-----FPVCTVVEYECRPGYRREPLSPKL---TCLQNLK-W---- 151
Qy 133 AAVPEGLFILFLHRRCFTLTQDEVQWCDHSSLPQSTPEIKHGPASASQVAGTKDMHHYTW 192

```
Db 152 -----STAVEFCKKSC-PNPGEIRN---GQIDVPG-----G 179
QY 193 LIFIFIENF-----LRQSLNS---VTQAGVQWENLGSLOPLRPGFKLFSCPSLLSSWD 242
Db 180 ILFGATISFCNTGKLFGSTSSFCLSGSSVQWSD-----PLPECREIY-CPA----- 227
QY 243 YRRPPLANFFV-----FLVEMGFTMFARLILI-----SGP----- 273
Db 228 ---PQIDNGIIQGERDHVGRQSVTYACNKGFTMIGEHSIYCTVNNDGEWSGPPPECR 284
QY 274 -----CDLP-----ASASQAGITGVSHH----- 292
Db 285 GKSLTSKVPPTVQKPTTVNVPTTEVSP7SQKTTTKTTTENAQATRSTPVSRTTKHFETT 344
QY 293 -----ARLIFNCLFEMESHVTOAGVQWPNLGSLOPLRPGFKLFSCLSLPSSW 341
Db 345 PNKSGTTSYTTRLISG-----SRPVTQAGMRWCDRSSLQSRTPGPKRSFHFLSPSSW 397
QY 342 DY 343
Db 398 YY 399
```

Search completed: September 8, 2005, 10:45:33
Job time : 72 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 8, 2005, 06:54:01 ; Search time 376 Seconds
(without alignments)
510.717 Million cell updates/sec

Title: US-10-092-934-10
Perfect score: 2034
Sequence: 1 MEFSLLPLRLCNGAISAHRR.....FIRGGVSPYLXSGWQTFDLR 375

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt 03:*

1: uniprot_sprot:*

2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2034	100.0	375	2	O60448
2	595	29.3	587	1	ALUJ2_HUMAN
3	568	27.9	587	1	ALUJ3_HUMAN
4	549.5	27.0	591	1	ALUJ8_HUMAN
5	498.5	24.5	591	1	ALUJ1_HUMAN
6	389	19.1	585	1	ALUJ5_HUMAN
7	377.5	18.6	158	2	Q8NAL9
8	357.5	17.6	593	1	ALUJ6_HUMAN
9	356.5	17.5	603	1	ALUJ4_HUMAN
10	349	17.2	157	2	Q8N845
11	337.5	16.6	164	2	O6ZTS0
12	332	16.3	176	2	O6ZRN2
13	331	16.3	170	2	O6ZV14
14	330.5	16.2	593	1	ALUJ7_HUMAN
15	325	16.0	143	2	Q6ZSJ7
16	320	15.7	195	2	O6ZUW2
17	316	15.5	152	2	Q8NX85
18	303.5	14.9	239	2	O6ZWA9
19	303	14.9	124	2	Q6ZVF6
20	301	14.8	156	2	Q8N266
21	295	14.5	193	2	O6ZVX8
22	294.5	14.5	148	2	O6ZUC5
23	290	14.3	123	2	O6ZV58
24	287.5	14.1	166	2	O6ZUC0
25	287	14.1	180	2	Q8N7M7
26	286.5	14.1	130	2	O6ZSV2
27	286.5	14.1	131	2	O6ZTD3
28	283	13.9	125	2	O6ZTE1
29	281.5	13.8	147	2	O6ZUA2
30	280	13.8	165	2	O6ZTF6
31	279.5	13.7	132	2	O6ZS28

RESULT 1									
O60448	ID	O60448	PRELIMINARY;	PRT;	375	AA.			
AC	O60448;								
DT	01-AUG-1998	(TrEMBLrel. 07, Created)							
DT	01-AUG-1998	(TrEMBLrel. 07, Last sequence update)							
DE	01-JUN-2003	(TrEMBLrel. 24, Last annotation update)							
DE	Neuronal thread protein AD7c-NTP.								
OS	Homo sapiens (Human).								
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;								
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.								
OX	NCBI_TaxID=9606;								
RN	[1]								
RP	SEQUENCE FROM N.A.								
RC	TISSUE=Neuronal;								
RX	MEDLINE=98064067; PubMed=9399956;								
RA	Monte S.M., Ghanbari K., Frey W.H., Beheshti I., Averbach P.,								
RA	Hauser S.L., Ghanbari H.A., Wands J.R.;								
RT	"Characterization of the AD7C-NTP cDNA expression in Alzheimer's								
RT	disease and measurement of a 41-kD protein in cerebrospinal fluid."								
RL	J. Clin. Invest. 100:3093-3104(1997).								
RN	[2]								
RP	SEQUENCE FROM N.A.								
RC	TISSUE=Neuronal;								
RA	de la Monte S.M., Wands J.R.;								
RL	Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.								
DR	EMBL; AF010144; AAC08737.1; ..								
DR	GO; GO:0005615; C:extracellular space; TAS.								
DR	GO; GO:0016021; C:integral to membrane; TAS.								
DR	GO; GO:0006915; P:apoptosis; TAS.								
DR	GO; GO:0007417; P:central nervous system development; TAS.								
SQ	SEQUENCE 375 AA; 41720 MW; 955443950ASBPED CRC64;								
Query Match 100.0%; Score 2034; DB 2; Length 375;									
Best Local Similarity 100.0%; Pred. NO. 1.9e-171;									
Matches 375; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
Qy	1	MEFSLLPLRLCNGAISAHRRALPGSSDSPASPSVAGITGMCTHARLLYFLVEMEF	60						
Db	1	MEFSLLPLRLCNGAISAHRRALPGSSDSPASPSVAGITGMCTHARLLYFLVEMEF	60						
Qy	61	LHVQAGLEPTSDPSVSASQARVYTHHARLLCLANFCGRNVSLMCPSPWSPKLQST	120						
Db	61	LHVQAGLEPTSDPSVSASQARVYTHHARLLCLANFCGRNVSLMCPSPWSPKLQST	120						
Qy	121	CLSIPKCDWYRRRAVPGFLFLFLRHRCPTLTQDEVQCDHSSLPQSTPIKHPASASQ	180						
Db	121	CLSIPKCDWYRRRAVPGFLFLFLRHRCPTLTQDEVQCDHSSLPQSTPIKHPASASQ	180						
Qy	181	VAGTKDMHYTWLIFIFIFNFRQSLNSVTQAGVQWRNLGSLQPLPGFKLFSCPSLSS	240						
Db	181	VAGTKDMHYTWLIFIFIFNFRQSLNSVTQAGVQWRNLGSLQPLPGFKLFSCPSLSS	240						
Qy	241	WDYRRPRLANFFVFLVEMGFTWPARLLISGPCDLPASASQSGITGVSHHARLLIFNFC	300						

Q6zwd5 homo sapien
Q6zwd4 homo sapien
Q8n8c2 homo sapien
Q6zuko0 homo sapien
Q9h387 homo sapien
Q6zul3 homo sapien
Q6zge8 homo sapien
Q6z897 homo sapien
Q6zyt3 homo sapien
Q6zuw8 homo sapien
Q6zqr8 homo sapien
Q6z853 homo sapien
Q6zt71 homo sapien
Q6zsu7 homo sapien

Db 241 WYRRPRLANFVFLVEMGFTMFARLILISGCDLPASASQAGITGVSHHARLIENFC 300
 Qy 301 LFEMESHVSQTQAGVQWPNLGSLOPLPGLKRFSCLSLPSWDYGHLPHPANFCIFIRGG 360
 Db 301 LFEMESHVSQTQAGVQWPNLGSLOPLPGLKRFSCLSLPSWDYGHLPHPANFCIFIRGG 360
 Qy 361 VSPYLSGWSQTPDLR 375
 Db 361 VSPYLSGWSQTPDLR 375

RESULT 2

ALU2 HUMAN
 ID ALU2 HUMAN STANDARD; PRT; 587 AA.
 AC P39189;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 25-OCT-2004 (Rel. 45, Last annotation update)
 DE Alu subfamily SB sequence contamination warning entry.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=95021758; PubMed=7935834; DOI=10.1038/371752a0;
 RA Claverie J.-M., Makalowski W.;
 RT "Alu alert."
 RL Nature 371:752-752(1994).
 RN [2]
 RP CONCEPT.
 RX MEDLINE=92241891; PubMed=1572661;
 RA Claverie J.-M.;
 RT "Identifying coding exons by similarity search: alu-derived and other
 potentially misleading protein sequences.";
 RL Genomics 12:838-841(1992).
 RN [3]
 RP ALU FAMILIES CLASSIFICATION.
 RX MEDLINE=88333009; PubMed=3138422;
 RA Quentin Y.;
 RT "The Alu family developed through successive waves of fixation closely
 connected with primate lineage history.";
 RL J. Mol. Evol. 27:194-202(1988).
 RN [4]
 RP ALU FAMILIES CLASSIFICATION.
 RX MEDLINE=91178815; PubMed=1706781;
 RA Jurka J., Milosavljevic A.;
 RT "Reconstruction and analysis of human Alu genes.";
 RL J. Mol. Evol. 32:105-121(1991).
 CC -!- MISCELLANEOUS: Various analyses indicate that Alu repeats fall
 into 8 subfamilies (Ref.3 and Ref.4). Therefore, 8 Alu warning
 consensus sequences have been constituted that contain all six
 frames conceptual translations of each of these classes of Alu
 repeats.
 CC -!- MISCELLANEOUS: Isolated 'X' indicates the presence of a stop
 codon, 'XXX' is used to separate the various translation phases.
 CC -!- CAUTION: This Alu entry is provided in order to avoid the further
 pollution of protein sequence databases with Alu-derived amino
 acid sequences.
 CC -!- CAUTION: Alu repetitive sequences are interspersed in human and
 primate genomes with an average spacing of 4 kb. Some of them are
 actively transcribed by pol III. Normal transcripts may contain
 Alu-derived sequences in 5' or 3' untranslated regions. However,
 cDNA libraries also contain partial and/or rearranged cDNAs
 ligated with Alu-derived sequence in any orientation. Although Alu
 elements (especially situated on the complementary strand) have a
 great potential to create additional/alternative exons,
 consideration should be given to the possibility that the presence
 of an Alu in an open reading frame may have resulted from a
 cloning artifact or may be due to misinterpretation of sequencing
 data. This point has been overlooked on several occasions, with
 the consequence of erroneous Alu-derived amino acid sequences

CC being reported.
 CC -!- CAUTION: Any significant similarity of a putative protein sequence
 with an Alu-translated entry must be taken as a warning that a
 part of Alu repeat may have been artifactually included in the
 coding nucleotide sequence.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (see <http://www.ebi.ac.uk/announcements/>
 or send an email to license@ebi.ac.uk).
 CC -----
 CC EMBL; U14568; -; NOT_ANNOTATED_CDS.
 DR Hypothetical protein.
 KW DOMAIN 1 96 Frame-1.
 FT DOMAIN 100 194 Frame-2.
 FT DOMAIN 198 292 Frame-3.
 FT DOMAIN 296 391 Frame-4.
 FT DOMAIN 395 489 Frame-5.
 FT DOMAIN 493 587 Frame-6.
 SQ SEQUENCE 587 AA; 63703 MW; 3EAB3E3E3929203 CRC64;
 Query Match 29.3%; Score 595; DB 1; Length 587;
 Best Local Similarity 41.5%; Pred. No. 5.7e-44;
 Matches 153; Conservative 14; Mismatches 96; Indels 106; Gaps 8;
 Qy 6 LLPRLECNAGISAHRNLRPGSSDPSASPVAGITGCTHARLILYFFLVEMEFHVQ 65
 Db 304 LSPRLCNGAISAHCKLRPGSRHSPASASRVAGTGTGHHARLI-FVFLVETGFRVSQ 362
 Qy 66 AGLELPTSDPSVSASQSAARYRTGHHARLCLANFCGRNRYSLMCPSPKQSTCLSLP 125
 Db 363 DGLDLLT-----SXSAR-----LGLP 378
 Qy 126 KCDYRRAAVPGLFILFRLHRCPTLTQDEVQVCDHSSLOPSTPEIKHPASASQVAGTK 185
 Db 379 KCDYRREPPRPAXXFF-----XDGVSIC-----RPGWSAVARSRLTASAS--- 421
 Qy 186 DMHHYTWLIFIFINFLRSLNSVTQAGVQWPNLGSLOPLPGLKRFSCLSLPSWDYGHLPHPANFCIFIRGGVSPYL 245
 Db 422 -----RVHAILLPQPEXIGLOAPATP----- 444
 Qy 246 PPRLANFVFLVEMGFTMFARLILISGCDLPASASQAGITGVSHHARLIENFCLEME 305
 Db 445 -----GXFLYFXRRGFTVTLARMVSIKPRDPPASASQAGITGVSHHARXXX---FFETE 497
 Qy 306 SHSVTQAGVQWPNLGSLOPLPGLKRFSCLSLPSWDYGHLPHPANFCIFIRGGVSPYL 365
 Db 498 SRVAQAGVQWRDLGSLQAPPGFTPFSCLSLPSWDYRRPPRPPANFCIFSRDGVSPCX 557
 Qy 366 SGMSQTPDL 374
 Db 558 PGWSRSPDL 566
 RESULT 3
 ALU3 HUMAN
 ID ALU3 HUMAN STANDARD; PRT; 587 AA.
 AC P39150;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Alu subfamily SBI sequence contamination warning entry.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=95021758; PubMed=7935834; DOI=10.1038/371752a0;
 RA Claverie J.-M., Makalowski W.;

RT "Alu alert."; Nature 371:752-752(1994).
 RL (2)
 RN
 RP CONCEPT.
 RX MEDLINE=92241891; PubMed=1572661;
 RA Claverie J.-M.;
 RT "Identifying coding exons by similarity search: alu-derived and other
 RT potentially misleading protein sequences."; Genomics 12:838-841(1992).
 RL (3)
 RN
 RP ALU FAMILIES CLASSIFICATION.
 RX MEDLINE=88333009; PubMed=3138422;
 RA Quentin Y.;
 RT "The Alu family developed through successive waves of fixation closely
 RT connected with primate lineage history."; J. Mol. Evol. 27:194-202(1988).
 RL (4)
 RN
 RP ALU FAMILIES CLASSIFICATION.
 RX MEDLINE=91178815; PubMed=1706781;
 RA Jurka J., Milosavljevic A.;
 RT "Reconstruction and analysis of human Alu genes."; J. Mol. Evol. 32:105-121(1991).
 RL
 CC -1- MISCELLANEOUS: Various analyses indicate that Alu repeats fall
 CC into 8 subfamilies (Ref.3 and Ref.4). Therefore, 8 Alu warning
 CC consensus sequences have been constituted that contain all six
 CC frames conceptual translations of each of these classes of Alu
 CC repeats.
 CC
 CC -1- MISCELLANEOUS: Isolated 'X' indicates the presence of a stop
 CC codon, 'XXX' is used to separate the various translation phases.
 CC
 CC -1- CAUTION: This Alu entry is provided in order to avoid the further
 CC pollution of protein sequence databases with Alu-derived amino
 CC acid sequences.
 CC
 CC -1- CAUTION: Alu repetitive sequences are interspersed in human and
 CC primate genomes with an average spacing of 4 kb. Some of them are
 CC actively transcribed by pol III. Normal transcripts may contain
 CC Alu-derived sequences in 5' or 3' untranslated regions. However,
 CC cDNA libraries also contain partial and/or rearranged cDNAs
 CC ligated with Alu-derived sequence in any orientation. Although Alu
 CC elements (especially situated on the complementary strand) have a
 CC great potential to create additional/alternative exons,
 CC consideration should be given to the possibility that the presence
 CC of an Alu in an open reading frame may have resulted from a
 CC cloning artifact or may be due to misinterpretation of sequencing
 CC data. This point has been overlooked on several occasions, with
 CC the consequence of erroneous Alu-derived amino acid sequences
 CC being reported.
 CC
 CC -1- CAUTION: Any significant similarity of a putative protein sequence
 CC with an Alu-translated entry must be taken as a warning that a
 CC part of Alu repeat may have been artifactually included in the
 CC coding nucleotide sequence.
 CC
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; U14569; -, NOT_ANNOTATED_CDS.
 DR Hypothetical protein.
 KW
 FT DOMAIN 1 96 Frame-1.
 FT DOMAIN 100 194 Frame-2.
 FT DOMAIN 198 292 Frame-3.
 FT DOMAIN 296 391 Frame-4.
 FT DOMAIN 395 489 Frame-5.
 FT DOMAIN 493 587 Frame-6.
 SQ SEQUENCE 587 AA; 63573 MW; 85C4155726DEF235 CRC64;
 Query Match 27.9%; Score 568; DB 1; Length 587;
 Best Local Similarity 41.5%; Pred. No. 1.4e-41;
 Matches 153; Conservative 15; Mismatches 95; Indels 106; Gaps 10;

QY 6 LLPRLECNAGISAHRNLRPLPGSSDSDSPASASPVAGITGCTHARILLYFFLVMEFLHVQ 65
 DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
 304 LSPRLCEGSGISAHCKRLPGSRHSPPASASQVAGTTGARHYARLI-FVFLVETGEHFRS 362
 QY 66 AGLELPTSDPPSVASASQARYRTGHARLCLANFCGNNRVSLMCPSPSPKOSTCISLP 125
 DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
 363 DGJDLIT-----SASAR-----LGLP 378
 QY 126 KCWDYRRAAVPGLFILFFLHRCPTLTQDEVOMCDHSSLOPSTPEIKHPPASASQVAGTK 185
 DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
 379 KCWDYRREP RPAXXXFF-----XDGVSIC-----RPGWSA--VAGSR 414
 QY 186 DMHYTWLIFIFINFLROSLNSVTOAGVOMRNLGSLQPLPPGFKLFSCSLSSSWDYRR 245
 DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
 415 -----LTASSASRVHA-----ILLPQPKXGLQAPATTP----- 444
 QY 246 PPRLANFFVFLVEMGTMTFARLILISGCDLPASASQAGITGVSHHARLIFNCLFEME 305
 DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
 445 -----GXFLYFXRXRGFTVLGAMVSI8XPRDPPASASQAGITGVSHHARXXX---PFETE 497
 QY 306 SHSVTOAGVOMRNLGSLQPLPPGFKLFSCSLSSWDYGHLPHPANFCIFIRGGVSPYL 365
 DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
 498 SRVAQGVOMRDLGSLQAPPPGFTPFSCSLSSWDYRRPPLRPANFCIFSRDGVSPFX 557
 QY 366 SGWSQTPDL 374
 DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
 558 PGWSRSPDL 566
 RESULT 4
 ID ALU8_HUMAN STANDARD; PRT; 591 AA.
 AC P39135;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 25-OCT-2004 (Rel. 45, Last annotation update)
 DE Alu subfamily SX sequence contamination warning entry.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=95021759; PubMed=7935834; DOI=10.1038/371752a0;
 RA Claverie J.-M., Makalowski W.;
 RL "Alu alert."; Nature 371:752-752(1994).
 RN [2]
 RP CONCEPT.
 RX MEDLINE=92241891; PubMed=1572661;
 RA Claverie J.-M.;
 RT "Identifying coding exons by similarity search: alu-derived and other
 RT potentially misleading protein sequences."; Genomics 12:838-841(1992).
 RL [3]
 RN ALU FAMILIES CLASSIFICATION.
 RX MEDLINE=88333009; PubMed=3138422;
 RA Quentin Y.;
 RT "The Alu family developed through successive waves of fixation closely
 RT connected with primate lineage history."; J. Mol. Evol. 27:194-202(1988).
 RL [4]
 RN ALU FAMILIES CLASSIFICATION.
 RX MEDLINE=91178815; PubMed=1706781;
 RA Jurka J., Milosavljevic A.;
 RT "Reconstruction and analysis of human Alu genes."; J. Mol. Evol. 32:105-121(1991).
 CC -1- MISCELLANEOUS: Various analyses indicate that Alu repeats fall
 CC into 8 subfamilies (Ref.3 and Ref.4). Therefore, 8 Alu warning
 CC consensus sequences have been constituted that contain all six
 CC frames conceptual translations of each of these classes of Alu
 CC repeats.
 CC
 CC -1- MISCELLANEOUS: Isolated 'X' indicates the presence of a stop

```

CC      codon, 'XXX' is used to separate the various translation phases.
CC      -!- CAUTION: This Alu entry is provided in order to avoid the further
CC      pollution of protein sequence databases with Alu-derived amino
CC      acid sequences.
CC      -!- CAUTION: Alu repetitive sequences are interspersed in human and
CC      primate genomes with an average spacing of 4 kb. Some of them are
CC      actively transcribed by pol III. Normal transcripts may contain
CC      Alu-derived sequences in 5' or 3' untranslated regions. However,
CC      cDNA libraries also contain partial and/or rearranged cDNAs
CC      ligated with Alu-derived sequence in any orientation. Although Alu
CC      elements (especially situated on the complementary strand) have a
CC      great potential to create additional/alternative exons,
CC      consideration should be given to the possibility that the presence
CC      of an Alu in an open reading frame may have resulted from a
CC      cloning artifact or may be due to misinterpretation of sequencing
CC      data. This point has been overlooked on several occasions, with
CC      the consequence of erroneous Alu-derived amino acid sequences
CC      being reported.
CC      -!- CAUTION: Any significant similarity of a putative protein sequence
CC      with an Alu-translated entry must be taken as a warning that a
CC      part of Alu repeat may have been artifactually included in the
CC      coding nucleotide sequence.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; U14574; -; NOT ANNOTATED_CDS.
KW      Hypothetical protein.
FT      DOMAIN 1 96 Frame-1.
FT      DOMAIN 100 195 Frame-2.
FT      DOMAIN 199 294 Frame-3.
FT      DOMAIN 298 393 Frame-4.
FT      DOMAIN 397 492 Frame-5.
FT      DOMAIN 496 591 Frame-6.
SQ      SEQUENCE 591 AA; 64395 MW; AC8154AD8A6BB280 CRC64;

Query Match      27.0%; Score 549.5; DB 1; Length 591;
Best Local Similarity 38.9%; Pred. No. 6.1e-40;
Matches 150; Conservative 15; Mismatches 84; Indels 137; Gaps 11;

QY      6 LLPLREKNGAISAHNRLPGSSDSPASASPVAGITGCTHARLLTYFLVEMFLHVQ 65
Db      306 LSPRLCEGSAISAHNRLPGSSDSPASASPVAGITGARHARLLI-FVLVETGFHHVQ 364

QY      66 AGLELPTSD--PSVSASQSAHYRTGTHARLCLANFCGRNVRSLMCPSSW-----P 114
Db      365 AGLELLTSGDLLPKVLGLQXATAPG---LRPXXKFFXGDSVLCRPGMSAVARSRLTAT 420

QY      115 ELKQSTCLSLPKCWD---YRRAVPGLFLLFLLRHRCPTLTQDEVQWCHSSLPSTPE 170
Db      421 SASRVAQAILLPQPPXELGLQARATPGXFLFYXRRGFTMLAR-----LVNSNXPQ 471

QY      171 IKHPPASASQVAGTKDMHHYTWLFIIFLFLNLRQSLNSVTQAGVQWRNLGSLQLPPGFK 230
Db      472 V-----IF----- 474

QY      231 LFPSCPLLSSWDYRR-PPRLANFFVFLVEMGFTMEARLLISGPCDLPASASQSGITGV 289
Db      475 -----LPKCWDYREPR-----PASARXXX----- 495

QY      290 SHHARLIENFLCFEMESHVSVTQAGVQWRNLGSLQLPPLGPKRFXCLSLPSSWDYGHLDPPH 349
Db      496 -----FFETESRSVAQGVQWRDLGSLQPPPPGKFRFXCLSLPSSWDYRRAPPR 544

QY      350 PANFCIFIRGGVSPYLSGWSQTPDLR 375
Db      545 PANFCIFSRDGVSPCWPNGWSRTPDLR 570

```

```

RESULT 5
ALU1_HUMAN
ID      ALU1_HUMAN STANDARD; PRT; 591 AA.
AC      P39188;
DT      01-FEB-1995 (Rel. 31, Created)
DT      01-FEB-1995 (Rel. 31, Last sequence update)
DT      25-OCT-2004 (Rel. 45, Last annotation update)
DE      Alu subfamily J sequence contamination warning entry.
OS      Homo sapiens (Human).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX      NCBI_TaxId=9606;
[1]
SEQUENCE FROM N.A.
RP      MEDLINE=95021758; PubMed=7935834; DOI=10.1038/371752a0;
RA      Claverie J.-M., Makalowski W.;
RT      "Alu alert.";
RL      Nature 371:752-752 (1994).
[2]
RP      CONCEPT.
RX      MEDLINE=92241891; PubMed=1572661;
RA      Claverie J.-M.;
RT      "Identifying coding exons by similarity search: alu-derived and other
RT      potentially misleading protein sequences.";
RL      Genomics 12:838-841 (1992).
[3]
RP      ALU FAMILIES CLASSIFICATION.
RX      MEDLINE=88333009; PubMed=3138422;
RA      Quentin Y.;
RT      "The Alu family developed through successive waves of fixation closely
RT      connected with primate lineage history.";
RL      J. Mol. Evol. 27:194-202 (1988).
[4]
RP      ALU FAMILIES CLASSIFICATION.
RX      MEDLINE=91178815; PubMed=1706781;
RA      Jurka J., Milosavljevic A.;
RT      "Reconstruction and analysis of human Alu genes.";
RL      J. Mol. Evol. 32:105-121 (1991).
CC      -!- MISCELLANEOUS: Various analyses indicate that Alu repeats fall
CC      into 8 subfamilies (Ref.3 and Ref.4). Therefore, 8 Alu warning
CC      consensus sequences have been constituted that contain all six
CC      frames conceptual translations of each of these classes of Alu
CC      repeats.
CC      -!- MISCELLANEOUS: Isolated 'X' indicates the presence of a stop
CC      codon, 'XXX' is used to separate the various translation phases.
CC      -!- CAUTION: This Alu entry is provided in order to avoid the further
CC      pollution of protein sequence databases with Alu-derived amino
CC      acid sequences.
CC      -!- CAUTION: Alu repetitive sequences are interspersed in human and
CC      primate genomes with an average spacing of 4 kb. Some of them are
CC      actively transcribed by pol III. Normal transcripts may contain
CC      Alu-derived sequences in 5' or 3' untranslated regions. However,
CC      cDNA libraries also contain partial and/or rearranged cDNAs
CC      ligated with Alu-derived sequence in any orientation. Although Alu
CC      elements (especially situated on the complementary strand) have a
CC      great potential to create additional/alternative exons,
CC      consideration should be given to the possibility that the presence
CC      of an Alu in an open reading frame may have resulted from a
CC      cloning artifact or may be due to misinterpretation of sequencing
CC      data. This point has been overlooked on several occasions, with
CC      the consequence of erroneous Alu-derived amino acid sequences
CC      being reported.
CC      -!- CAUTION: Any significant similarity of a putative protein sequence
CC      with an Alu-translated entry must be taken as a warning that a
CC      part of Alu repeat may have been artifactually included in the
CC      coding nucleotide sequence.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial

```


CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL; U14567; -; NOT ANNOTATED_CDS.

DR Hypothetical protein_96 Frame-1.
 FT DOMAIN 1 96
 FT DOMAIN 100 195 Frame-2.
 FT DOMAIN 199 294 Frame-3.
 FT DOMAIN 298 393 Frame-4.
 FT DOMAIN 397 492 Frame-5.
 FT DOMAIN 496 591 Frame-6.
 SQ SEQUENCE 591 AA; 63790 MW; 665D395735519D95 CRC64;

Query Match 24.5%; Score 498.5; DB 1; Length 591;
 Best Local Similarity 38.6%; Pred. No. 2e-35;
 Matches 144; Conservative 15; Mismatches 103; Indels 111; Gaps 11;

Qy 6 LLPRLENGAISAHRNLRPGSSDPASAPVAGITGCTHARLILYFFVEMFLHVQ 65
 Db 306 LSPRLCSGAIATACSLDLPGSSDPASAPVAGITGCTHARLILYFFVEMFLHVQ 364

Qy 66 AGLEPTSDPSVASOSARYRTG--HHARLCLANFCGRNVRSLMCPSPKQSTCLS 123
 Db 365 AGLELSSDPASASAGI-TVSHR--XXXFFXDRVSLCRPGNSA----- 411

Qy 124 LPKCDYRRAAVPGLFLFLFRLHRCPTLTQDEVQWCDHSSLOPSTPPIKHPASASQVAG 183
 Db 412 -----VARSRLTAASRAQ-----AIIQQPPE----- 435

Qy 184 TKDMHVTWLIPIFIFNLRQSLNSVTQAGVQWRNLSGLQPLPGFKLFCPSLLSSWDY 243
 Db 436 -----XLGLQAR-----ATTPG----- 447

Qy 244 RRPRLANFFVELVEMGFTMFARLILIS-GPCDLPASASOSAGITGVSHHARLIFNCLF 302
 Db 448 -----XFLYFLXRRGFTMLPRVSNMAQVILPPRPK---VLGLQATAPGXKXFF 497

Qy 303 EMBSHSVTQAGVQWRNLSGLQPLPGFKLFCPSLLSSWDYGHLPHPANFCIFIRGVS 362
 Db 498 ETGSRVAQAGVQWRNLSGLQPLPGFKLFCPSLLSSWDYRRAPPRANFCIFCRDVS 557

Qy 363 PYLSGWSQTPDLR 375
 Db 558 LCCPGWSRTPELK 570

RESULT 6

ALU5 HUMAN STANDARD; PRT; 585 AA.
 AC P39192;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 25-OCT-2004 (Rel. 45, Last annotation update)
 DE Alu subfamily SC sequence contamination warning entry.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=95021758; PubMed=7935834; DOI=10.1038/371752a0;
 RA Claverie J.-M., Makalowski W.;
 RT "Alu alert."
 RL Nature 371:752-752(1994).
 RN [2]
 RP CONCEPT.
 RX MEDLINE=92241891; PubMed=1572661;
 RA Claverie J.-M.;
 RT "Identifying coding exons by similarity search: alu-derived and other
 potentially misleading protein sequences."
 RL Genomics 12:838-841(1992).
 RN [3]
 RP ALU FAMILIES CLASSIFICATION.

RX MEDLINE=88333009; PubMed=3138422;
 RA Quentin Y.;
 RT "The Alu family developed through successive waves of fixation closely
 connected with primate lineage history.";
 RL J. Mol. Evol. 27:194-202(1988).
 RN [4]
 RP ALU FAMILIES CLASSIFICATION.
 RX MEDLINE=91178815; PubMed=1706781;
 RA Jurka J., Milosavljevic A.;
 RT "Reconstruction and analysis of human Alu genes.";
 RL J. Mol. Evol. 32:105-121(1991).
 CC -I- MISCELLANEOUS: Various analyses indicate that Alu repeats fall
 into 8 subfamilies (Ref.3 and Ref.4). Therefore, 8 Alu warning
 consensus sequences have been constituted that contain all six
 frames conceptual translations of each of these classes of Alu
 repeats.

CC -I- MISCELLANEOUS: Isolated 'X' indicates the presence of a stop
 codon, 'XX' is used to separate the various translation phases.
 CC -I- CAUTION: This Alu entry is provided in order to avoid the further
 pollution of protein sequence databases with Alu-derived amino
 acid sequences.

CC -I- CAUTION: Alu repetitive sequences are interspersed in human and
 primate genomes with an average spacing of 4 kb. Some of them are
 actively transcribed by pol III. Normal transcripts may contain
 Alu-derived sequences in 5' or 3' untranslated regions. However,
 cDNA libraries also contain partial and/or rearranged cDNAs
 ligated with Alu-derived sequence in any orientation. Although Alu
 elements (especially situated on the complementary strand) have a
 great potential to create additional/alternative exons,
 consideration should be given to the possibility that the presence
 of an Alu in an open reading frame may have resulted from a
 cloning artifact or may be due to misinterpretation of sequencing
 data. This point has been overlooked on several occasions, with
 the consequence of erroneous Alu-derived amino acid sequences
 being reported.

CC -I- CAUTION: Any significant similarity of a putative protein sequence
 with an Alu-translated entry must be taken as a warning that a
 part of Alu repeat may have been artifactually included in the
 coding nucleotide sequence.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

CC EMBL; U14571; -; NOT ANNOTATED_CDS.

KW Hypothetical protein.
 FT DOMAIN 1 95 Frame-1.
 FT DOMAIN 99 193 Frame-2.
 FT DOMAIN 197 291 Frame-3.
 FT DOMAIN 295 389 Frame-4.
 FT DOMAIN 393 487 Frame-5.
 FT DOMAIN 491 585 Frame-6.
 SQ SEQUENCE 585 AA; 63957 MW; 46B8C4F493650A7 CRC64;

Query Match 19.1%; Score 389; DB 1; Length 585;
 Best Local Similarity 33.9%; Pred. No. 9.5e-26;
 Matches 130; Conservative 23; Mismatches 100; Indels 130; Gaps 17;

Qy 17 SAHRNLRPGSSDPASAPVAGITGCTHARLILYFFVEMFLHVQAGLEPTSDDP 76
 Db 189 SVSKXXXPGVAHACNPSTLGGRGGRITRSR-----DRDPGQHG-----ETP 232

Qy 77 SVSASQARYTGHARLCLANFCGR-----NRV-----SLMC-PSWSPKQST 120
 Db 233 SLKIKIQLAGRG--ARL-XSQLGLRLQENRLNPGGGCGSEPRSRHCTPAW----RQSE 285

Qy 121 CLSLPKCDYRRAAVPGLFLFLFRLHRCPTLTQDEVQWCDHSSLOPSTPPIKHPASASQ 180
 Db 286 TPSQKK-----XXXFFLRSL-----ALSPGW-----SA 309

```
QY 181 VAGTKDMHHYTWLIFIFIFNFIQSLNSVTOAGVQWRNLGSLQPLPP---GFKLFSCPSL 237
|||
Db 310 VA-----RSLTATSASRVQ-----AALLQPPXELGLQ----- 338
|
QY 238 LSSWYDRPPRIANFVFLVNGFTMFARLILISGCDLPASASQASAGITGVSHHARLIF 297
|||
Db 339 -----ARATTPSXFLYFXRRGFTMLARMVSIISPRDPPASASQASAGITGVSHRXXX 392
|
QY 298 NF-----CLPEMESHVSTOAGVQWRNLGSLQPLPPGLKRFSCSLSPSSWDYGHLPHPA 351
|||
Db 393 FFXDGVSLC-----RQAGVQWRDLGSLQPLPPPGFKRFSCSLSPSSWDYRRPRPA 443
|
QY 352 NFCIFIRGVSPLYLGGWGTDPDL 374
|||
Db 444 NFCIFSRDGVSPFCWGNRSIDL 466
|
RESULT 7
QBNAL9 PRELIMINARY; PRT; 158 AA.
AC QBNAL9;
DT 01-OCT-2002 (TRENBLrel. 22, Created)
DT 01-OCT-2002 (TRENBLrel. 22, Last sequence update)
DT 01-OCT-2002 (TRENBLrel. 22, Last annotation update)
DE Hypothetical protein FLJ35131.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RX PubMed=14702039; DOI=10.1038/ngl285;
RA Ota T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R.,
RA Wakamatsu A., Hayashi K., Sato H., Nagai K., Kimura K., Makita H.,
RA Sekine M., Ohyaashi M., Nishi T., Shibahara T., Tanaka T., Ishii S.,
RA Yamamoto K., Saito K., Kawai Y., Isono Y., Nakamura Y., Nagahari K.,
RA Murakami K., Yasuda T., Iwayanagi T., Wagatsuma M., Shiratori A.,
RA Sudo H., Hosoiri T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,
RA Takahashi M., Kanda K., Yokoi T., Furuya T., Kikkawa E., Omura Y.,
RA Abe K., Kamihara K., Katsuta N., Sato K., Tanikawa M., Yamazaki M.,
RA Ninomiya K., Ishibashi T., Yamashita H., Murakawa K., Fujimori K.,
RA Tanai H., Kimata M., Watanabe M., Hirao K., Chiba Y., Ishida S.,
RA Ono Y., Takiguchi S., Watanabe S., Yosida M., Hotuta T., Kusano J.,
RA Kanehori K., Takahashi-Fujii A., Hara H., Tanase T., Nomura Y.,
RA Togliya S., Komai F., Hara R., Takeuchi K., Arita M., Imose N.,
RA Musashino K., Yuuki H., Oshima A., Sasaki N., Aotsuka S.,
RA Yoshihara Y., Matsunawa H., Ichihara T., Shiohata N., Sano S.,
RA Moriya S., Momiyama H., Satoh N., Takami S., Terashima Y., Suzuki O.,
RA Nakagawa S., Senoh A., Mizoguchi H., Goto Y., Shinizu F., Wakebe H.,
RA Hishigaki H., Watanabe T., Sugiyama A., Takemoto M., Kawakami B.,
RA Yamazaki M., Watanabe K., Kumagai A., Itakura S., Fukuzumi Y.,
RA Fujimori Y., Komiyama M., Tashiro H., Tanigami A., Fujiwara T.,
RA Ono T., Yamada K., Fujii Y., Ozaki K., Hirao M., Ohmori Y.,
RA Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,
RA Okitani R., Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T.,
RA Matsumura K., Nakajima Y., Mizuno T., Morinaga M., Sasaki M.,
RA Togashi T., Oyama M., Hata H., Watanabe M., Komatsu T.,
RA Mizushima-Sugano J., Satoh T., Shirai Y., Takahashi Y., Nakagawa K.,
RA Okumura K., Nagase T., Nomura N., Kikuchi H., Masuho Y., Yamashita R.,
RA Nakai K., Yada T., Nakamura Y., Ohara O., Isogai T., Sugano S.;
RT "Complete sequencing and characterization of 21,243 full-length human
RT cDNAs";
RL Nat. Genet. 36:40-45(2004).
DR EMBL; AK092450; BAC03893.1; -.
SQ SEQUENCE 158 AA; 17414 MW; E151503015F2FE34 CRC64;
Query Match 18.6%; Score 377.5; DB 2; Length 158;
Best Local Similarity 53.3%; Pred. No. 2.2e-25;
Matches 81; Conservative 7; Mismatches 21; Indels 43; Gaps 3;
QY 207 NSVTQAGVQWRNLGSLQPLPPGFKLFSCPSLSSWDYRR-PPRLANFFVFLVNGFTMFA 265
```

```
Db 4 HSVTQAGVQWRNLGSLQPLPPGFKLFSCPSLSSWDYRRVPPPLANFCIF----- 53
|||
QY 266 RLILISGCDLPASASQASAGITGVSHHARLIFNFCLEPMESHVSTOAGVQWRNLGSLQPL 325
|||
Db 54 -----FSP-FFEKESLSVTQAGVQWRDLGSLQAA 81
|
QY 326 PPGKRFSCSLSPSSWDYGHLPHPANFCIFI 357
|||
Db 82 PPGFTFSCSLSPSSWYRRPPPCPANFFVL 113
|
RESULT 8
ALU6_HUMAN STANDARD; PRT; 593 AA.
AC P39193;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DB Alu subfamily SP sequence contamination warning entry.
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=92241891; PubMed=1572661;
RA Claverie J.-M.;
RT "Identifying coding exons by similarity search: alu-derived and other
RT potentially misleading protein sequences.";
RL Genomics 12:838-841(1992).
RN [3]
RP ALU FAMILIES CLASSIFICATION.
RX MEDLINE=88333009; PubMed=3138422;
RA Quentin Y.;
RT "The Alu family developed through successive waves of fixation closely
RT connected with primate lineage history.";
RL J. Mol. Evol. 27:194-202(1988).
RN [4]
RP ALU FAMILIES CLASSIFICATION.
RX MEDLINE=91178815; PubMed=1706781;
RA Jurka J., Milosavljevic A.;
RT "Reconstruction and analysis of human Alu genes.";
RL J. Mol. Evol. 32:105-121(1991).
CC -1- MISCELLANEOUS: Various analyses indicate that Alu repeats fall
CC into 8 subfamilies (Ref.3 and Ref.4). Therefore, 8 Alu warning
CC consensus sequences have been constituted that contain all six
CC frames conceptual translations of each of these classes of Alu
CC repeats.
CC -1- MISCELLANEOUS: Isolated 'X' indicates the presence of a stop
CC codon, 'XXX' is used to separate the various translation phases.
CC -1- CAUTION: This Alu entry is provided in order to avoid the further
CC pollution of protein sequence databases with Alu-derived amino
CC acid sequences.
CC -1- CAUTION: Alu repetitive sequences are interspersed in human and
CC primate genomes with an average spacing of 4 kb. Some of them are
CC actively transcribed by pol III. Normal transcripts may contain
CC Alu-derived sequences in 5' or 3' untranslated regions. However,
CC cDNA libraries also contain partial and/or rearranged cDNAs
CC ligated with Alu-derived sequence in any orientation. Although Alu
CC elements (especially situated on the complementary strand) have a
CC great potential to create additional/alternative exons,
CC consideration should be given to the possibility that the presence
CC of an Alu in an open reading frame may have resulted from a
CC cloning artifact or may be due to misinterpretation of sequencing
CC data. This point has been overlooked on several occasions, with
CC the consequence of erroneous Alu-derived amino acid sequences
```

being reported.
 -!- CAUTION: Any significant similarity of a putative protein sequence with an Alu-translated entry must be taken as a warning that a part of Alu repeat may have been artifactually included in the coding nucleotide sequence.

 This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

 EMBL; U14572; -; NOT_ANNOTATED_CDS.
 ProDom; PD003738; GIDA; 1.
 Hypothetical protein.
 DOMAIN 1 97 Frame-1.
 DOMAIN 101 196 Frame-2.
 DOMAIN 200 295 Frame-3.
 DOMAIN 299 395 Frame-4.
 DOMAIN 399 494 Frame-5.
 DOMAIN 498 593 Frame-6.
 SEQUENCE 593 AA; 64603 MW; 136EF344AACD12A2 CRC64;
 Query Match 17.6%; Score 357.5; DB 1; Length 593;
 Best Local Similarity 48.2%; Pred. No. 5.9e-23;
 Matches 95; Conservative 6; Mismatches 47; Indels 49; Gaps 8;
 ;
 QY 4 SLLPLRLCEGCAISAHNRLPGSSDSPASPVAGITGCMTHARLLYFLVEMELHV 63
 DQ 405 SLLPLRLCEGCAISAHNRLPGSSDSPASPVAGITGCMTHARLLYFLVETGFLHV 463
 QY 64 GQAGLELPTSDPVSASQSARYRTG--HHARLCLANFCGRRNRLVSLMCPSPWS----- 113
 DQ 464 GQAGLELPTSDPVSASQSAGI-TGVSHRAR---XXXFFETFRSCPGNSAMARSLT 519
 QY 114 -----PELKOSTCLSLPKCWDYRRAAVPGILFLFLRHRCPTLLQDEVQW 158
 DQ 520 ATSASRVQAILLPQPEX-----LGLQAC-----ATTPGXFLYEXRRGFSMLVR--- 564
 QY 159 CDHSSLOPSTPEIKHPP 175
 DQ 565 ----LVNSNRPOQVIRPP 577
 RESULT 9
 ID _ALU4_HUMAN STANDARD; PRT; 603 AA.
 AC P39191;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 25-OCT-2004 (Rel. 45, Last annotation update)
 DE Alu subfamily SB2 sequence contamination warning entry.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=95021758; PubMed=7935834; DOI=10.1038/371752a0;
 RA Claverie J.-M.; Makalowski W.;
 RT "Alu alert."
 RL Nature 371:752-752 (1994).
 RN [2]
 RP CONCEPT.
 RX MEDLINE=92241891; PubMed=1572661;
 RA Claverie J.-M.;
 RT "Identifying coding exons by similarity search: alu-derived and other potentially misleading protein sequences";
 RL Genomics 12:838-841 (1992).
 RN [3]
 RP ALU FAMILIES CLASSIFICATION.

RX MEDLINE=88333009; PubMed=3138422;
 RA Quentin Y.;
 RT "The Alu family developed through successive waves of fixation closely connected with primate lineage history.";
 RL J. Mol. Evol. 27:194-202 (1988).
 RN [4]
 RP ALU FAMILIES CLASSIFICATION.
 RX MEDLINE=91178815; PubMed=1706781;
 RA Jurka J., Milosavljevic A.;
 RT "Reconstruction and analysis of human Alu genes.";
 RL J. Mol. Evol. 32:105-121 (1991).
 CC -!- MISCELLANEOUS: Various analyses indicate that Alu repeats fall into 8 subfamilies (Ref.3 and Ref.4). Therefore, 8 Alu warning consensus sequences have been constituted that contain all six frames conceptual translations of each of these classes of Alu repeats.
 CC -!- MISCELLANEOUS: Isolated 'X' indicates the presence of a stop codon, 'XXX' is used to separate the various translation phases.
 CC -!- CAUTION: This Alu entry is provided in order to avoid the further pollution of protein sequence databases with Alu-derived amino acid sequences.
 CC -!- CAUTION: Alu repetitive sequences are interspersed in human and primate genomes with an average spacing of 4 kb. Some of them are actively transcribed by pol III. Normal transcripts may contain Alu-derived sequences in 5' or 3' untranslated regions. However, cDNA libraries also contain partial and/or rearranged cDNAs ligated with Alu-derived sequence in any orientation. Although Alu elements (especially situated on the complementary strand) have a great potential to create additional/alternative exons, consideration should be given to the possibility that the presence of an Alu in an open reading frame may have resulted from a cloning artifact or may be due to misinterpretation of sequencing data. This point has been overlooked on several occasions, with the consequence of erroneous Alu-derived amino acid sequences being reported.
 CC -!- CAUTION: Any significant similarity of a putative protein sequence with an Alu-translated entry must be taken as a warning that a part of Alu repeat may have been artifactually included in the coding nucleotide sequence.

 This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

 EMBL; U14570; -; NOT_ANNOTATED_CDS.
 KW Hypothetical protein.
 FT DOMAIN 1 98 Frame-1.
 FT DOMAIN 102 199 Frame-2.
 FT DOMAIN 202 300 Frame-3.
 FT DOMAIN 304 401 Frame-4.
 FT DOMAIN 405 502 Frame-5.
 FT DOMAIN 506 603 Frame-6.
 SQ SEQUENCE 603 AA; 65272 MW; B8AAD0AD46BEA114 CRC64;
 Query Match 17.5%; Score 356.5; DB 1; Length 603;
 Best Local Similarity 30.3%; Pred. No. 7.3e-23;
 Matches 112; Conservative 28; Mismatches 115; Indels 115; Gaps 14;
 QY 76 PVSASQSARYRTGHHARLCLANFCGRRNRLVSLMCPSPKOSTCLSLPKCWDYRRAAV 135
 DQ 156 PSYGGXGRWAXTREAEALVSRDCA---TAVRSAPWATERDSVS-----KXXXX 202
 QY 136 PGLFILFLRHRC-PTTQDEVQW-----DHSSLOPSTP----- 169
 DQ 203 PGA-----VAHACNPSTLGGGWMRSGRDHPGXQGETPSLLKIQKISRARWRAPVVP 257
 QY 170 -----ELKHP-----PASASQVAGTKDMHHYTWLFIFFIENFLR 203
 DQ 258 ATGEAEAGEWREPQKRSLOXAEIAPLQSAVRPGRQSETPSQKXXF-----FLR 307
 ALU FAMILIES CLASSIFICATION.


```

DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein FLJ46542.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Thymus;
RA Ninomiya K., Wagatsuma M., Kanda K., Kondo H., Yokoi T., Kodaira H.,
RA Furuya T., Takahashi M., Kikkawa E., Omura Y., Abe K., Kamihara K.,
RA Katsuta N., Sato K., Tanikawa M., Yamazaki M., Sugiyama T., Irie R.,
RA Otsuki T., Sato H., Ota T., Wakamatsu A., Ishii S., Yamamoto J.,
RA Isono Y., Kawai-Hio Y., Saito K., Nishikawa T., Kimura K.,
RA Yamashita H., Matsuo K., Nakamura Y., Sekine M., Kikuchi H.,
RA Murakawa K., Kanehori K., Takahashi-Fujii A., Oshima A., Sugiyama A.,
RA Kawakami B., Suzuki Y., Sugano S., Nagahara K., Masuho Y., Nagai K.,
RA Isogai T.;
RL Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK128399; BAC87422.1; -.
SQ SEQUENCE 176 AA; 19060 MW; 98D1CAC1940AD8FA CRC64;

Query Match 16.3%; Score 332; DB 2; Length 176;
Best Local Similarity 47.9%; Pred. No. 2.7e-21;
Matches 90; Conservative 11; Mismatches 51; Indels 36; Gaps 8;

Qy 208 SVTQAGVQWRNLGSLQPLPPGKLFSCPSLLSSWDYRRPPPLAN-----FF 253
Dy 5 SVSQAQGVQWRD-----PPAHCIL---HLSSSD---SPASARVAIVISGTCHHAQLIF 51

Qy 254 VFLVEMGFTMPA-----RLILISGCDLPASASOSAGITGVSHARLIFNP-CLFEMESH 308
Dy 52 VFLVETGFCYVQAGRLRLTSSDP---PTSASLSAGIAGLSHRARPFCLFLCVFTVSHS 108

Qy 309 VTQAGVQWPNLGSQPLPPGKLFSCPSLLSSWDYRHPPHPAN-FCIFIRGGVSPYLSG 367
Dy 109 VTQGMQWCDHSSLSQSRPPGLKRSPLSSWDYRQASPLHANYFFFFIRGKASLCCPG 168

Qy 368 WSQTPDLR 375
Dy 169 WSQTPELK 176

RESULT 13
Q6ZVI4 PRELIMINARY; PRT; 170 AA.
AC Q6ZVI4;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DE Hypothetical protein FLJ42545.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Cerebellum;
RA Ninomiya K., Wagatsuma M., Kanda K., Kondo H., Yokoi T., Kodaira H.,
RA Furuya T., Takahashi M., Kikkawa E., Omura Y., Abe K., Kamihara K.,
RA Katsuta N., Sato K., Tanikawa M., Yamazaki M., Sugiyama T., Irie R.,
RA Otsuki T., Sato H., Wakamatsu A., Ishii S., Yamamoto J., Isono Y.,
RA Kawai-Hio Y., Saito K., Nishikawa T., Kimura K., Yamashita H.,
RA Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Murakawa K.,
RA Kanehori K., Takahashi-Fujii A., Oshima A., Sugiyama A., Kawakami B.,
RA Suzuki Y., Sugano S., Nagahara K., Masuho Y., Nagai K., Isogai T.;
RL Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK124536; BAC85877.1; -.
SQ SEQUENCE 170 AA; 19351 MW; A019980274CEA078 CRC64;

Query Match 16.3%; Score 331; DB 2; Length 170;
Best Local Similarity 42.0%; Pred. No. 3.1e-21;
Matches 71; Conservative 16; Mismatches 39; Indels 44; Gaps 3;

```

```

Qy 251 NFFVFLVEMGFTMPARL-----ILISGCDLPASASOSAGITG-----VSH 291
Dy 2 NRLFYKSGQGLALLPRLKSCGAIHAHCNFFELGSSNFPDILASERAGTATFGTVVLIRLSN 61
Qy 292 HARLIENF-----CLFEMESHVTOAGVQWPNLGSLOPLP 326
Dy 62 HIAMLDWFWRKQTIWSTRNLNHHLVSCISFIIIFETESHVTOAGVQWPNLGSLOPLP 121

Qy 327 PGLKRFSCLSLPSWDYHGLPHHPANFCIFIRGGVSPYLSGWSQTPDLR 375
Dy 122 PGKRFSCSLTPTSDYMQMPPCLANFCIFSRDGVSPYWPWGWSRTPDCR 170

RESULT 14
ALU7 HUMAN
ID ALU7 HUMAN STANDARD; PRT; 593 AA.
AC P39194;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Alu subfamily SQ sequence contamination warning entry.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Claverie J.-M., Makalowski W.;
RT "Alu alert.";
RL Nature 371:752-752 (1994).
RN [2]
RP CONCEPT.
RA MEDLINE=92241891; PubMed=1572661;
RX Claverie J.-M.;
RT "Identifying coding exons by similarity search: alu-derived and other
potentially misleading protein sequences.";
RL Genomics 12:838-841 (1992).
RN [3]
RP ALU FAMILIES CLASSIFICATION.
RX MEDLINE=88333009; PubMed=3138422;
RA Quentin Y.;
RT "The Alu family developed through successive waves of fixation closely
connected with primate lineage history.";
RL J. Mol. Evol. 27:194-202 (1988).
RN [4]
RP ALU FAMILIES CLASSIFICATION.
RX MEDLINE=91178815; PubMed=1706781;
RA Jurka J., Milosavljevic A.;
RT "Reconstruction and analysis of human Alu genes.";
RL J. Mol. Evol. 32:105-121 (1991).
CC -1- MISCELLANEOUS: Various analyses indicate that Alu repeats fall
into 8 subfamilies (Ref.3 and Ref.4). Therefore, 8 Alu warning
consensus sequences have been constituted that contain all six
frames conceptual translations of each of these classes of Alu
repeats.
CC -1- MISCELLANEOUS: Isolated 'x' indicates the presence of a stop
codon, 'xxx' is used to separate the various translation phases.
CC -1- CAUTION: This Alu entry is provided in order to avoid the further
pollution of protein sequence databases with Alu-derived amino
acid sequences.
CC -1- CAUTION: Alu repetitive sequences are interspersed in human and
primate genomes with an average spacing of 4 kb. Some of them are
actively transcribed by pol III. Normal transcripts may contain
Alu-derived sequences in 5' or 3' untranslated regions. However,
cDNA libraries also contain partial and/or rearranged cDNAs
ligated with Alu-derived sequence in any orientation. Although Alu
elements (especially situated on the complementary strand) have a
great potential to create additional/alternative exons,
consideration should be given to the possibility that the presence
of an Alu in an open reading frame may have resulted from a
cloning artifact or may be due to misinterpretation of sequencing

```


GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 8, 2005, 09:57:54 ; Search time 101 Seconds
(without alignments)
277.162 Million cell updates/sec

Title: US-10-092-934-10

Perfect score: 2034

Sequence: 1 MEFSLLPLRLECGAISAHR.....FIRGGVSPVLSGWSQTPDLR 375

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:*

- 1: /cgn2_6/ptodata/1/iaa/5A COMB.pdp.*
- 2: /cgn2_6/ptodata/1/iaa/5B COMB.pdp.*
- 3: /cgn2_6/ptodata/1/iaa/6A COMB.pdp.*
- 4: /cgn2_6/ptodata/1/iaa/6B COMB.pdp.*
- 5: /cgn2_6/ptodata/1/iaa/PCTUS COMB.pdp.*
- 6: /cgn2_6/ptodata/1/iaa/backfiles1.pdp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2034	100.0	375	2	US-08-454-557C-121
2	2034	100.0	375	2	US-08-340-426D-121
3	2034	100.0	375	2	US-08-450-673C-121
4	2034	100.0	375	4	US-09-872-968-2
5	1415.5	69.6	397	5	PCT-US95-17111A-121
6	267.5	13.2	96	4	US-09-513-999C-6065
7	259.5	12.8	144	4	US-09-513-999C-6953
8	256	12.6	108	4	US-09-513-999C-7878
9	254.5	12.5	239	4	US-09-800-729-193
10	254.5	12.5	310	4	US-09-800-729-190
11	253	12.4	118	4	US-09-663-600A-114
12	251	12.3	776	4	US-10-020-079-24
13	251	12.3	789	4	US-10-020-079-22
14	251	12.3	863	4	US-10-020-079-32
15	251	12.3	884	4	US-10-020-079-30
16	251	12.3	889	4	US-10-020-079-20
17	251	12.3	895	4	US-10-020-079-18
18	251	12.3	976	4	US-10-020-079-28
19	251	12.3	982	4	US-10-020-079-26
20	250	12.3	97	4	US-09-513-999C-4770
21	250	12.3	132	4	US-09-636-215-573
22	250	12.3	132	4	US-09-685-166A-573
23	250	12.3	132	4	US-09-679-426-573
24	250	12.3	132	4	US-09-759-143-573
25	250	12.3	132	4	US-09-651-236-573
26	250	12.3	135	4	US-09-685-166A-884
27	250	12.3	135	4	US-09-679-426-884

28	250	12.3	135	4	US-09-759-143-884	Sequence 884, App
29	247.5	12.2	112	4	US-09-513-999C-7870	Sequence 7870, App
30	246	12.1	119	4	US-09-513-999C-7867	Sequence 7867, App
31	245	12.0	102	4	US-09-621-976-6112	Sequence 6112, App
32	242.5	11.9	121	4	US-09-513-999C-7874	Sequence 7874, App
33	241	11.8	91	4	US-09-621-976-5929	Sequence 5929, App
34	239	11.3	84	4	US-09-621-976-7167	Sequence 7167, App
35	226	11.1	61	4	US-09-513-999C-4581	Sequence 4581, App
36	224.5	11.0	396	4	US-09-949-016-6783	Sequence 6783, App
37	224.5	11.0	442	4	US-09-949-016-9762	Sequence 9762, App
38	223.5	11.0	87	4	US-09-621-976-5968	Sequence 5968, App
39	219	10.8	76	4	US-09-621-976-6338	Sequence 6338, App
40	218	10.7	1079	3	US-09-058-489-22	Sequence 22, Appl
41	217.5	10.7	87	4	US-09-205-258-342	Sequence 342, App
42	216	10.6	169	4	US-09-863-600A-208	Sequence 208, App
43	215.5	10.6	108	4	US-09-513-999C-4433	Sequence 4433, App
44	214.5	10.5	83	4	US-09-621-976-5396	Sequence 5396, App
45	211.5	10.4	103	4	US-09-513-999C-5327	Sequence 5327, App

ALIGNMENTS

RESULT 1

US-08-454-557C-121
; Sequence 121, Application US/08454557C
; Patent No. 5830670
; GENERAL INFORMATION:
; APPLICANT: de la Monte, Suzanne
; APPLICANT: Wands, Jack R.
; TITLE OF INVENTION: Neural Thread Protein Gene Expression and Detection
; TITLE OF INVENTION: of Alzheimer's Disease
; NUMBER OF SEQUENCES: 121
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
; STREET: 1100 New York Avenue, Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/454,557C
; FILING DATE: 30-MAY-1995
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Ludwig, Steven R.
; REGISTRATION NUMBER: 36,203
; REFERENCE/DOCKET NUMBER: 0609,3840003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 121:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 375 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-454-557C-121

Query Match		100.0%;	Score 2034;	DB 2;	Length 375;
Best Local Similarity		100.0%;	Pred. No. 6.2e-215;		
Matches 375;		Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	MEFSLLPLRLECGAISAHRNLRPLGSSDSPASPSVAGITGCTHARLLYFFELVEMEF	60		
Db	1	MEFSLLPLRLECGAISAHRNLRPLGSSDSPASPSVAGITGCTHARLLYFFELVEMEF	60		
Qy	61	LHVQAGLELPTSDPVSASQSAARYTGHARHLCIANFCGRNRVSLMCPSPWSPDLQST	120		


```
Db 1 LHVGAGLEPTSDPSVSASQARYRTGHARLCLANFCGRNVRSLMCPWSPELKQST 120
Qy 121 CLSLPKCDWYRRAAVPGLFILFHLRHRCPTLTQDEVQWCDHSSLOPSTPEIKHPASASQ 180
Db 121 CLSLPKCDWYRRAAVPGLFILFHLRHRCPTLTQDEVQWCDHSSLOPSTPEIKHPASASQ 180
Qy 181 VAGTKDMHHTWLIIFIFNFRQSLNSVTQAGVQWNRNLGSLQPLPGFKLFSCPSLLSS 240
Db 181 VAGTKDMHHTWLIIFIFNFRQSLNSVTQAGVQWNRNLGSLQPLPGFKLFSCPSLLSS 240
Qy 241 WDYRRPRLANFFVFLVEMGFTMFARLILISGPCDLPASASQAGITGVSHHARLIENFC 300
Db 241 WDYRRPRLANFFVFLVEMGFTMFARLILISGPCDLPASASQAGITGVSHHARLIENFC 300
Qy 301 LFEMESHVTOAGVQWPNLGSLOPLPGKRFKFCSLSPSSWDYGHLPHPANFCIFIRGG 360
Db 301 LFEMESHVTOAGVQWPNLGSLOPLPGKRFKFCSLSPSSWDYGHLPHPANFCIFIRGG 360
Qy 361 VSPYLSGWSQTPDLR 375
Db 361 VSPYLSGWSQTPDLR 375
```

RESULT 2

```
US-08-340-426D-121
; Sequence 121, Application US/08340426D
; Patent No. 5948634
; GENERAL INFORMATION:
; APPLICANT: de la Monte, Suzanne
; APPLICANT: Wanda, Jack R.
; TITLE OF INVENTION: Neural Thread Protein Gene Expression and Detection
; TITLE OF INVENTION: of Alzheimer's Disease
; NUMBER OF SEQUENCES: 121
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
; STREET: 1100 New York Avenue, Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/340.426D
; FILING DATE: 14-NOV-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Ludwig, Steven R.
; REGISTRATION NUMBER: 36,203
; REFERENCE/DOCKET NUMBER: 0609.3840002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 121:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 375 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-340-426D-121
```

```
Query Match 100.0%; Score 2034; DB 2; Length 375;
Best Local Similarity 100.0%; Pred. No. 6.2e-215;
Matches 375; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MFSLLLPRLCNGAISAHRNLRLPGSSDSPASASPVAGITGMCTHARLILYFFLVEMEF 60
Db 1 MFSLLLPRLCNGAISAHRNLRLPGSSDSPASASPVAGITGMCTHARLILYFFLVEMEF 60
```

```
Qy 61 LHVGAGLEPTSDPSVSASQARYRTGHARLCLANFCGRNVRSLMCPWSPELKQST 120
Db 61 LHVGAGLEPTSDPSVSASQARYRTGHARLCLANFCGRNVRSLMCPWSPELKQST 120
Qy 121 CLSLPKCDWYRRAAVPGLFILFHLRHRCPTLTQDEVQWCDHSSLOPSTPEIKHPASASQ 180
Db 121 CLSLPKCDWYRRAAVPGLFILFHLRHRCPTLTQDEVQWCDHSSLOPSTPEIKHPASASQ 180
Qy 181 VAGTKDMHHTWLIIFIFNFRQSLNSVTQAGVQWNRNLGSLQPLPGFKLFSCPSLLSS 240
Db 181 VAGTKDMHHTWLIIFIFNFRQSLNSVTQAGVQWNRNLGSLQPLPGFKLFSCPSLLSS 240
Qy 241 WDYRRPRLANFFVFLVEMGFTMFARLILISGPCDLPASASQAGITGVSHHARLIENFC 300
Db 241 WDYRRPRLANFFVFLVEMGFTMFARLILISGPCDLPASASQAGITGVSHHARLIENFC 300
Qy 301 LFEMESHVTOAGVQWPNLGSLOPLPGKRFKFCSLSPSSWDYGHLPHPANFCIFIRGG 360
Db 301 LFEMESHVTOAGVQWPNLGSLOPLPGKRFKFCSLSPSSWDYGHLPHPANFCIFIRGG 360
Qy 361 VSPYLSGWSQTPDLR 375
Db 361 VSPYLSGWSQTPDLR 375
```

RESULT 3

```
US-08-450-673C-121
; Sequence 121, Application US/08450673C
; Patent No. 5948888
; GENERAL INFORMATION:
; APPLICANT: de la Monte, Suzanne
; APPLICANT: Wanda, Jack R.
; TITLE OF INVENTION: Neural Thread Protein Gene Expression and Detection
; TITLE OF INVENTION: of Alzheimer's Disease
; NUMBER OF SEQUENCES: 121
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
; STREET: 1100 New York Avenue, Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/450.673C
; FILING DATE: 30-MAY-1995
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Ludwig, Steven R.
; REGISTRATION NUMBER: 36,203
; REFERENCE/DOCKET NUMBER: 0609.3840004
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 121:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 375 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-450-673C-121
```

```
Query Match 100.0%; Score 2034; DB 2; Length 375;
Best Local Similarity 100.0%; Pred. No. 6.2e-215;
Matches 375; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MFSLLLPRLCNGAISAHRNLRLPGSSDSPASASPVAGITGMCTHARLILYFFLVEMEF 60
Db 1 MFSLLLPRLCNGAISAHRNLRLPGSSDSPASASPVAGITGMCTHARLILYFFLVEMEF 60
```

QY 61 LHVQAGLELPTSDPSVASQARVYTGHHARLCLANFCGRNRVSLMCPSPKQST 120
DB 61 LHVQAGLELPTSDPSVASQARVYTGHHARLCLANFCGRNRVSLMCPSPKQST 120
QY 121 CLSLPKCWDYRRAAVPGLFILFRLHRCPTLTQDEVQWCDHSSLOPSTPEIKHPPASQ 180
DB 121 CLSLPKCWDYRRAAVPGLFILFRLHRCPTLTQDEVQWCDHSSLOPSTPEIKHPPASQ 180
QY 181 VAGTKDMHYTWLIFIFINFLRQSLNSVTQAGVQWRNLSLOPLPPGKLFSCPSLLSS 240
DB 181 VAGTKDMHYTWLIFIFINFLRQSLNSVTQAGVQWRNLSLOPLPPGKLFSCPSLLSS 240
QY 241 WDYRRPRLANFFVFLVEMGFTMFARLILISGCDLPASASQAGITGVSHARLIFNFC 300
DB 241 WDYRRPRLANFFVFLVEMGFTMFARLILISGCDLPASASQAGITGVSHARLIFNFC 300
QY 301 LFEMESHVTOAGVQWPNLGSLOPLPGLKRFCSLPSWDYGHLPHPANFCIFIRGG 360
DB 301 LFEMESHVTOAGVQWPNLGSLOPLPGLKRFCSLPSWDYGHLPHPANFCIFIRGG 360
QY 361 VSPYLSGWSQTPDLR 375
DB 361 VSPYLSGWSQTPDLR 375
RESULT 4
US-09-872-968-2
; Sequence 2, Application US/09872968
; Patent No. 6770797
; GENERAL INFORMATION:
; APPLICANT: Wands, Jack R
; APPLICANT: de la Monte, Suzanne M
; TITLE OF INVENTION: Inhibition of Neurodegeneration
; FILE REFERENCE: 21486-047
; CURRENT APPLICATION NUMBER: US/09/872, 968
; CURRENT FILING DATE: 2001-06-01
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: Patentin ver. 2.1
; SEQ ID NO 2
; TYPE: PRT
; LENGTH: 375
; ORGANISM: Homo sapiens
US-09-872-968-2

Query Match 100.0%; Score 2034; DB 4; Length 375;
Best Local Similarity 100.0%; Pred. No. 6.2e-215;
Matches 375; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MEFSLLLPRLECNCAISAHNRNLRPGSSDSPASASPVAGITGCTHARLILYFVLEMEF 60
DB 1 MEFSLLLPRLECNCAISAHNRNLRPGSSDSPASASPVAGITGCTHARLILYFVLEMEF 60
QY 61 LHVQAGLELPTSDPSVASQARVYTGHHARLCLANFCGRNRVSLMCPSPKQST 120
DB 61 LHVQAGLELPTSDPSVASQARVYTGHHARLCLANFCGRNRVSLMCPSPKQST 120
QY 121 CLSLPKCWDYRRAAVPGLFILFRLHRCPTLTQDEVQWCDHSSLOPSTPEIKHPPASQ 180
DB 121 CLSLPKCWDYRRAAVPGLFILFRLHRCPTLTQDEVQWCDHSSLOPSTPEIKHPPASQ 180
QY 181 VAGTKDMHYTWLIFIFINFLRQSLNSVTQAGVQWRNLSLOPLPPGKLFSCPSLLSS 240
DB 181 VAGTKDMHYTWLIFIFINFLRQSLNSVTQAGVQWRNLSLOPLPPGKLFSCPSLLSS 240
QY 241 WDYRRPRLANFFVFLVEMGFTMFARLILISGCDLPASASQAGITGVSHARLIFNFC 300
DB 241 WDYRRPRLANFFVFLVEMGFTMFARLILISGCDLPASASQAGITGVSHARLIFNFC 300
QY 301 LFEMESHVTOAGVQWPNLGSLOPLPGLKRFCSLPSWDYGHLPHPANFCIFIRGG 360
DB 301 LFEMESHVTOAGVQWPNLGSLOPLPGLKRFCSLPSWDYGHLPHPANFCIFIRGG 360

QY 361 VSPYLSGWSQTPDLR 375
DB 361 VSPYLSGWSQTPDLR 375
RESULT 5
PCT-US95-17111A-121
; Sequence 121, Application PC/TUS9517111A
; GENERAL INFORMATION:
; APPLICANT: de la Monte, Suzanne
; APPLICANT: Wands, Jack R.
; TITLE OF INVENTION: Neural Thread Protein Gene Expression and
; TITLE OF INVENTION: Detection of Alzheimer's Disease
; NUMBER OF SEQUENCES: 121
; CORRESPONDENCE ADDRESS:
; ADDRESSES: Sterne, Kessler, Goldstein & Fox P.L.L.C.
; STREET: 1100 New York Avenue, Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/17111A
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/340,426
; FILING DATE: 14-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Ludwig, Steven R.
; REGISTRATION NUMBER: 36,203
; REFERENCE/DOCKET NUMBER: 0609.3840002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 121:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 397 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
PCT-US95-17111A-121
Query Match 69.6%; Score 1415.5; DB 5; Length 397;
Best Local Similarity 74.6%; Pred. No. 6.8e-147;
Matches 285; Conservative 10; Mismatches 44; Indels 43; Gaps 7;
QY 1 MEFSLLLPRLECNCAISAHNRNLRPGSSDSPASASPVAGITGCTHARLILYFVLEMEF 60
DB 1 MEFSLLLPRLECNCAISAHNRNLRPGSSDSPASASPVAGITGCTHARLILYFVLEMEF 60
QY 61 LHVQAGLELPTSDPSVASQARVYTGHHARLCLANFCGRNRVSLMCP 110
DB 61 LHVQAGLELPTSDPSVASQARVYTGHHARLCLANFCGRNRVSLMCP 110
QY 111 SWSPELKQSTCLSLPKCWDYRRAAVPGLFILFRLHRCPTLTQDEVQWCDHSSLOPSTPE 170
DB 111 SWSPELKQSTCLSLPKCWDYRRAAVPGLFILFRLHRCPTLTQDEVQWCDHSSLOPSTPE 170
QY 171 ----IKHPASASQVAGTKDMHYTWLIFIFINFLRQSLNSVTQAGVQWRNLSLOPLP 226
DB 171 SSILLQPPP----KVAGTKDMHYTWLIFIFINFLRQSLNSVTQAGVQWRNLSLOPLP 226
QY 227 PGFKLFSCPSLLSSWDYRRPRLANFFVFLVEMGFTMFARLILISGCDLPASASQAGITGVSHARLIFNFC 278
DB 227 PGFKLFSCPSLLSSWDYRRPRLANFFVFLVEMGFTMFARLILISGCDLPASASQAGITGVSHARLIFNFC 278
QY 279 SASQSAGITGVSHARLIFNFCIFEMESHVTOAGVQWPNLGSLOPLPPGKLFSCPSLLSS 338

```

FILE REFERENCE: 59 US2.REG
CURRENT APPLICATION NUMBER: US/09/513,999C
CURRENT FILING DATE: 2000-02-24
PRIOR APPLICATION NUMBER: US 60/122,487
PRIOR FILING DATE: 1999-02-26
NUMBER OF SEQ ID NOS: 36681
SOFTWARE: Patent.pm
SEQ ID NO 6953
LENGTH: 144
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: UNSURE
LOCATION: 9
OTHER INFORMATION: Xaa= * or Cys or Trp
FEATURE:
NAME/KEY: UNSURE
LOCATION: 10
OTHER INFORMATION: Xaa=Leu or Met or Val

```

```

1152: PK1
/ ORGANISM: Homo sapiens
/
/ FEATURE:
/ NAME/KEY: SIGNAL
/ LOCATION: -17...-1
/ OTHER INFORMATION: score 5.7
/ OTHER INFORMATION: seq VVCCSCSWFFLFC/FV
/
/ FEATURE:
/ NAME/KEY: UNSURE
/ LOCATION: 56
/ OTHER INFORMATION: Xaa= * or Cys or Gly
/
/ FEATURE:
/ NAME/KEY: UNSURE
/ LOCATION: 59
/ OTHER INFORMATION: Xaa=Arg or Ser
/
/ FEATURE:
/ NAME/KEY: UNSURE
/ LOCATION: 61
/ OTHER INFORMATION: Xaa=Asp or Asn
/
/ FEATURE:
/ NAME/KEY: UNSURE
/ LOCATION: 62
/ OTHER INFORMATION: Xaa=Ala or Glu or Gly
/
/ FEATURE:
/ NAME/KEY: UNSURE
/ LOCATION: 64
/ OTHER INFORMATION: Xaa= * or Leu
/

```

```

; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: 66
; OTHER INFORMATION: Xaa=Asn or Tyr
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: 67
; OTHER INFORMATION: Xaa= * or Tyr
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: 71
; OTHER INFORMATION: Xaa=Pro or Ser
US-09-513-999C-7878

Query Match      12.6%; Score 256; DB 4; Length 108;
Best Local Similarity 66.7%; Pred. No. 3.2e-20;
Matches 54; Conservative 3; Mismatches 20; Indels 4; Gaps 3;

Qy 297 FNFC-LFEMESHVTOAGVQVNPNGSLQPLPGKRFSCLSLPSSWDYGHLPHPAN--P 353
Db 14 FLFCFVEMESHVTOAGVQVNDLGLQALPPGFSPF-CLSLPSSWDYRCCLPPSSANFFF 72

Qy 354 CIFIRGGVSPYLSGWSQTPDL 374
Db 73 XIFXXXVXXPGXQSPDL 93

RESULT 9
US-09-800-729-193
; Sequence 193, Application US/09800729
; Patent No. 6605592
; GENERAL INFORMATION:
; APPLICANT: Ni et al.
; TITLE OF INVENTION: 32 Human secreted proteins
; FILE REFERENCE: P2044P1
; CURRENT APPLICATION NUMBER: US/09/800,729
; PRIOR FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: PCT/US00/26013
; PRIOR FILING DATE: 2000-09-22
; PRIOR APPLICATION NUMBER: 60/155,709
; PRIOR FILING DATE: 1999-09-24
; NUMBER OF SEQ ID NOS: 217
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 193
; LENGTH: 239
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-800-729-193

Query Match      12.5%; Score 254.5; DB 4; Length 239;
Best Local Similarity 64.2%; Pred. No. 1.5e-19;
Matches 61; Conservative 6; Mismatches 25; Indels 3; Gaps 3;

Qy 1 MEFSLL-PRLECNCAISAHNRNLPFGSSDSPASVPVAGITGCTHARLILYFVLVME 59
Db 120 LRWSLTSPRLCSCSAISAHNLRNLPFGSSNSPALASQVAGITGICHARQI-FVFLVETG 178

Qy 60 FLHVQAGLELPTSDPDSVSASQAR-YRTGHAR 93
Db 179 FCHVQAGLELLISGDSPPASAFQSGAGITGVSHR 213

RESULT 10
US-09-800-729-190
; Sequence 190, Application US/09800729
; Patent No. 6605592
; GENERAL INFORMATION:
; APPLICANT: Ni et al.
; TITLE OF INVENTION: 32 Human secreted proteins
; FILE REFERENCE: P2044P1
; CURRENT APPLICATION NUMBER: US/09/800,729
; CURRENT FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: PCT/US00/26013

```

```

; PRIOR FILING DATE: 2000-09-22
; PRIOR APPLICATION NUMBER: 60/155,709
; PRIOR FILING DATE: 1999-09-24
; NUMBER OF SEQ ID NOS: 217
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 190
; LENGTH: 310
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-800-729-190

Query Match      12.5%; Score 254.5; DB 4; Length 310;
Best Local Similarity 64.2%; Pred. No. 2.2e-19;
Matches 61; Conservative 6; Mismatches 25; Indels 3; Gaps 3;

Qy 1 MEFSLL-PRLECNCAISAHNRNLPFGSSDSPASVPVAGITGCTHARLILYFVLVME 59
Db 191 LRWSLTSPRLCSCSAISAHNLRNLPFGSSNSPALASQVAGITGICHARQI-FVFLVETG 249

Qy 60 FLHVQAGLELPTSDPDSVSASQAR-YRTGHAR 93
Db 250 FCHVQAGLELLISGDSPPASAFQSGAGITGVSHR 284

RESULT 11
US-09-663-600A-114
; Sequence 114, Application US/09663600A
; Patent No. 6573068
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, Jean-Baptiste
; APPLICANT: Duclert, Aymeric
; APPLICANT: Bougueleret, Lydie
; TITLE OF INVENTION: EXTENDED CDNAS FOR SECRETED PROTEINS
; FILE REFERENCE: 31.US3.CIP
; CURRENT APPLICATION NUMBER: US/09/663,600A
; CURRENT FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: 09/191,997
; PRIOR FILING DATE: 1998-11-13
; PRIOR APPLICATION NUMBER: 60/066,677
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/069,957
; PRIOR FILING DATE: 1997-12-17
; PRIOR APPLICATION NUMBER: 60/074,121
; PRIOR FILING DATE: 1998-02-09
; PRIOR APPLICATION NUMBER: 60/081,563
; PRIOR FILING DATE: 1998-04-13
; PRIOR APPLICATION NUMBER: 60/096,116
; PRIOR FILING DATE: 1998-08-10
; PRIOR APPLICATION NUMBER: 60/099,273
; PRIOR FILING DATE: 1998-09-04
; NUMBER OF SEQ ID NOS: 229
; SOFTWARE: Patent.pm
; SEQ ID NO 114
; LENGTH: 118
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: -83...-1
; NAME/KEY: UNSURE
; LOCATION: 28,32
; OTHER INFORMATION: Xaa = any one of the twenty amino acids
US-09-663-600A-114

Query Match      12.4%; Score 253; DB 4; Length 118;
Best Local Similarity 65.9%; Pred. No. 7.9e-20;
Matches 60; Conservative 3; Mismatches 24; Indels 4; Gaps 3;

Qy 204 QSLNSVTOAGVQVNRNLSQLPLPGFKLFSCPSLLSSWDYRRPRL-ANFFVFLVEMGFT 262
Db 5 QSFLVAQAGVQVNRHLSQLPLPPEFKGFCLSLPSSWDYRRPPCPAGFVFLVETGLH 64

Qy 263 MF--ARLILISGPCDLPASASQSGAGITGVSH 291

```

RESULT 13
US-10-020-079-22
; Sequence 22, Application US/10020079
; Patent No. 6579710
; GENERAL INFORMATION:
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Mathur, Brian
; APPLICANT: Friddle, Carl Johan
; TITLE OF INVENTION: No. 6579710el Human Kinases and Polynucleotides Encoding the Same
; FILE REFERENCE: LEX-0281-USA
; CURRENT APPLICATION NUMBER: US/10/020,079
; CURRENT FILING DATE: 2001-12-12
; PRIOR APPLICATION NUMBER: US 60/255,103
; PRIOR FILING DATE: 2000-12-12
; PRIOR APPLICATION NUMBER: US 60/289,422
; PRIOR FILING DATE: 2001-05-08
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: Fast-SEQ for Windows Version 4.0
; SEQ ID NO 22
; LENGTH: 789
; TYPE: PRT
; ORGANISM: homo sapiens
; FEATURE:
; NAME/KEY: VARIANT

```

Query Match      12.3%; Score 251; DB 4; Length 863;
Best Local Similarity 47.9%; Pred. No. 2.3e-18;
Matches 69; Conservative 9; Mismatches 38; Indels 28; Gaps 8;

QY      211 QAGVQWNLGSLQPLPPGKFLFCSPSLLSSWDYRR--PPRLANFEVFLVEMGFT-MFARL 267
          |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db       736 QAGVQWRDLGSLQPPPPRFKQFCSLPSRWDVTRHAPPPHAP-FVFLVETGFLHVEAGL 794
          |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::

QY      268 -LLISGPCDLPASASQAGITGVSHHARLIENFCLFEMESHSHSVTQAGVQWNLGSLQPLP 326
          ::|||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db       795 ELLTSG--DLPPASASQIAGITGVSHRAQP--EVCEFN-RKHTGQR----- 834
          ::|||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::

QY      327 PGLKRFCSLLPSSWDYGHLPHP 350
          ::|||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db       835 ---EQMVYAGSERAWSMRDLPGRP 855
          ::|||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::

RESULT 15
US-10-020-079-30
; Sequence 30, Application US/10020079
; Patent No. 6579710
; GENERAL INFORMATION:
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Mathuz, Brian

```

APPLICANT: Friddle, Carl Johan
TITLE OF INVENTION: No. 6579710e1 Human Kinases and Polynucleotides Encoding the Same
FILE REFERENCE: LEX-0281-USA
CURRENT APPLICATION NUMBER: US/10/020,079
CURRENT FILING DATE: 2001-12-12
PRIOR APPLICATION NUMBER: US 60/255,103
PRIOR FILING DATE: 2000-12-12
PRIOR APPLICATION NUMBER: US 60/289,422
PRIOR FILING DATE: 2001-05-08
NUMBER OF SEQ ID NOS: 40
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 30
LENGTH: 876
TYPE: PRT
ORGANISM: homo sapiens
FEATURE:
NAME/KEY: VARIANT
LOCATION: (1)...(876)
OTHER INFORMATION: Xaa = Any Amino Acid
US-10-020-079-30

Query Match 12.3%; Score 251; DB 4; Length 876;
Best Local Similarity 47.9%; Pred. No. 2.4e-18;
Matches 69; Conservative 9; Mismatches 38; Indels 28; Gaps 8;
Qy 211 QAGVQWRNLGSLQPLPFGKLFSCPSLLSSWDYRR--PPRLANFFVFLVEMGFT-MPEARL 267
Db 749 QAGVQWRDLGSLQPPPPRFRKQFCLSLPRSWDYRHAPPPHPAN-FVFLVETGFLHVEAGL 807
Qy 268 -ILISGCDLPASASQAGITGVSHARLIENFCLFEMESHVSVTQAGVQWRNLGSLQPLP 326
Db 808 ELLTSG--DLPASASQIAGITGVSHRAQP--EVCEFN-RKHTGQR----- 847
Qy 327 PGLKRFSCLSLPSSWDYGHLPHP 350
Db 848 ----EQMVCAGSERAWMRDLPGRP 868

Search completed: September 8, 2005, 10:47:21
Job time : 102 secs

THIS PAGE BLANK

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 8, 2005, 10:44:23 ; Search time 389 seconds
(without alignments)
379.650 Million cell updates/sec

Title: US-10-092-934-10

Perfect score: 2034

Sequence: 1 MEPSLLPLRLECNAGIAHR.....FIRGGVSPYLSGWSQPPDLR 375

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1774312 seqs, 393823214 residues

Total number of hits satisfying chosen parameters: 1774312

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

Published Applications AA:*

1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*

2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*

3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*

4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*

5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*

6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*

7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*

8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*

9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*

10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*

11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*

12: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*

13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*

14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*

15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*

16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep.*

17: /cgn2_6/ptodata/2/pubpaa/US10E_PUBCOMB.pep.*

18: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*

19: /cgn2_6/ptodata/2/pubpaa/US11A_PUBCOMB.pep.*

20: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep.*

21: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*

22: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2034	100.0	375	9	US-09-964-666-2
2	2034	100.0	375	9	US-09-964-412-2
3	2034	100.0	375	10	US-09-964-667-2
4	2034	100.0	375	10	US-09-872-968-2
5	2034	100.0	375	10	US-09-964-678A-2
6	2034	100.0	375	14	US-10-146-130-2
7	2034	100.0	375	14	US-10-092-934-10
8	2034	100.0	375	14	US-10-153-334-1
9	2034	100.0	375	14	US-10-198-069-1
10	2034	100.0	375	14	US-10-157-031-299
11	2034	100.0	375	14	US-10-198-070-1

12	2034	100.0	375	16	US-10-755-889-410	Sequence 410, App
13	2034	100.0	375	17	US-10-910-173-2	Sequence 2, Appl
14	734.5	36.1	361	9	US-09-995-494-107	Sequence 107, Appl
15	729	35.8	449	14	US-10-007-280A-140	Sequence 140, App
16	569	28.0	241	15	US-10-276-774-1834	Sequence 1834, Ap
17	517	25.4	213	15	US-10-296-115-911	Sequence 911, App
18	471	23.2	286	15	US-10-291-172-654	Sequence 654, App
19	471	23.2	286	15	US-10-221-278-654	Sequence 654, App
20	428.5	21.1	183	9	US-09-989-920-245	Sequence 245, App
21	407	20.0	361	15	US-10-276-774-1862	Sequence 1862, Ap
22	382.5	18.8	341	18	US-10-220-335-586	Sequence 586, App
23	382.5	18.8	673	14	US-10-157-031-291	Sequence 291, App
24	377.5	18.6	158	15	US-10-104-047-3047	Sequence 3047, Ap
25	356.5	17.5	603	16	US-10-408-765A-140	Sequence 140, App
26	353.5	17.4	217	14	US-10-017-161-1956	Sequence 1956, Ap
27	353.5	17.4	217	15	US-10-292-798-1604	Sequence 1604, Ap
28	349	17.2	157	15	US-10-108-260A-4272	Sequence 4272, Ap
29	336	16.5	148	15	US-10-276-774-1784	Sequence 1784, Ap
30	332	16.3	108	14	US-10-078-090-143	Sequence 143, App
31	324	15.9	189	15	US-10-104-047-3196	Sequence 3196, Ap
32	316	15.5	152	14	US-10-198-070-112	Sequence 112, App
33	311	15.3	74	15	US-10-276-774-1929	Sequence 1929, Ap
34	310	15.2	101	15	US-10-094-749-2181	Sequence 2181, Ap
35	309.5	15.2	107	15	US-10-276-774-1874	Sequence 1874, Ap
36	308	15.1	93	10	US-09-764-891-5337	Sequence 5337, Ap
37	308	15.1	93	14	US-10-205-428-486	Sequence 486, App
38	308	15.1	100	13	US-10-016-157A-187	Sequence 187, App
39	307	15.1	114	15	US-10-104-047-2423	Sequence 2423, Ap
40	305.5	15.0	137	15	US-10-276-774-1959	Sequence 1959, Ap
41	303	14.9	203	15	US-10-108-260A-2748	Sequence 2748, Ap
42	302.5	14.9	128	15	US-10-276-774-2194	Sequence 2194, Ap
43	301	14.8	156	15	US-10-104-047-2334	Sequence 2334, Ap
44	300.5	14.8	101	15	US-10-276-774-2026	Sequence 2026, Ap
45	300	14.7	76	15	US-10-276-774-2032	Sequence 2032, Ap

ALIGNMENTS

RESULT 1

US-09-964-666-2

; Sequence 2, Application US/09964666

; Patent No. US20020104108A1

; GENERAL INFORMATION:

; APPLICANT: de la Monte, Suzanne

; TITLE OF INVENTION: Transgenic Animals and Cell Lines for Screening Drugs Effective for the Treatment or Prevention of Alzheimer's Disease

; NUMBER OF SEQUENCES: 14

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Sterne, Kessler, Goldstein & Fox, P.L.L.C.

; STREET: 1100 New York Ave., Suite 600

; CITY: Washington

; STATE: DC

; COUNTRY: USA

; ZIP: 20005-3934

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/964,666

; FILING DATE: 28-Sep-2001

; CLASSIFICATION: <Unknown>

; ATTORNEY/AGENT INFORMATION:

; NAME: Esmond, Robert W.

; REGISTRATION NUMBER: 32,893

; REFERENCE/DOCKET NUMBER: 0609.4370000

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 202-371-2600

; TELEFAX: 202-371-2540

```
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 375 amino acids
;   TYPE: amino acid
;   TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-964-666-2

Query Match      100.0%; Score 2034; DB 9; Length 375;
Best Local Similarity 100.0%; Pred. No. 1e-184;
Matches 375; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MEFSLLLPRLECNCGAISAHNRNLRPGSSDSPASASPVAGITGMCTHARLILYFFLVEMEF 60
Db 1 MEFSLLLPRLECNCGAISAHNRNLRPGSSDSPASASPVAGITGMCTHARLILYFFLVEMEF 60

Qy 61 LHVGQAGLELPTSDPSVSASQASARYRTGHARLCLANFCGRNRVSLMCPSPWSPKQST 120
Db 61 LHVGQAGLELPTSDPSVSASQASARYRTGHARLCLANFCGRNRVSLMCPSPWSPKQST 120

Qy 121 CLSLPKCWDYRRAAVPGLFIFLFRHRCPTLTQDEVQWCDHSSLSQSPSTPEIKHPASASQ 180
Db 121 CLSLPKCWDYRRAAVPGLFIFLFRHRCPTLTQDEVQWCDHSSLSQSPSTPEIKHPASASQ 180

Qy 181 VAGTKDMHYTWLIFIFNFRQSLNSVTQAGVQWRNLSLQPLPGFKLFSCPSLSS 240
Db 181 VAGTKDMHYTWLIFIFNFRQSLNSVTQAGVQWRNLSLQPLPGFKLFSCPSLSS 240

Qy 241 WDYRRPRLANFFVFLVEMGTFMFARLILISGPCDLPASASQASAGITGVSHARLIENFC 300
Db 241 WDYRRPRLANFFVFLVEMGTFMFARLILISGPCDLPASASQASAGITGVSHARLIENFC 300

Qy 301 LFEMESHVTVQAGVQWPNLGLQPLPGKLRFSCLSLPSSWDYGHLPHPANFCIFIRGG 360
Db 301 LFEMESHVTVQAGVQWPNLGLQPLPGKLRFSCLSLPSSWDYGHLPHPANFCIFIRGG 360

Qy 361 VSPYLSGWSQTPDLR 375
Db 361 VSPYLSGWSQTPDLR 375

RESULT 2
US-09-964-412-2
; Sequence 2, Application US/09964412
; Patent No. US20020129391A1
; GENERAL INFORMATION:
; APPLICANT: de la Monte, Suzanne
; Wands, Jack R.
; TITLE OF INVENTION: Transgenic Animals and Cell Lines for
; Screening Drugs Effective for the Treatment or Prevention
; of Alzheimer's Disease
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox, P.L.L.C.
; STREET: 1100 New York Ave., Suite 600
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/964,412
; FILING DATE: 28-Sep-2001
; CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Esmond, Robert W.
; REGISTRATION NUMBER: 32,893
; REFERENCE/DOCKET NUMBER: 0609.4370000
```

```
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-2600
; TELEFAX: 202-371-2540
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 375 amino acids
;   TYPE: amino acid
;   TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-964-412-2

Query Match      100.0%; Score 2034; DB 9; Length 375;
Best Local Similarity 100.0%; Pred. No. 1e-184;
Matches 375; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MEFSLLLPRLECNCGAISAHNRNLRPGSSDSPASASPVAGITGMCTHARLILYFFLVEMEF 60
Db 1 MEFSLLLPRLECNCGAISAHNRNLRPGSSDSPASASPVAGITGMCTHARLILYFFLVEMEF 60

Qy 61 LHVGQAGLELPTSDPSVSASQASARYRTGHARLCLANFCGRNRVSLMCPSPWSPKQST 120
Db 61 LHVGQAGLELPTSDPSVSASQASARYRTGHARLCLANFCGRNRVSLMCPSPWSPKQST 120

Qy 121 CLSLPKCWDYRRAAVPGLFIFLFRHRCPTLTQDEVQWCDHSSLSQSPSTPEIKHPASASQ 180
Db 121 CLSLPKCWDYRRAAVPGLFIFLFRHRCPTLTQDEVQWCDHSSLSQSPSTPEIKHPASASQ 180

Qy 181 VAGTKDMHYTWLIFIFNFRQSLNSVTQAGVQWRNLSLQPLPGFKLFSCPSLSS 240
Db 181 VAGTKDMHYTWLIFIFNFRQSLNSVTQAGVQWRNLSLQPLPGFKLFSCPSLSS 240

Qy 241 WDYRRPRLANFFVFLVEMGTFMFARLILISGPCDLPASASQASAGITGVSHARLIENFC 300
Db 241 WDYRRPRLANFFVFLVEMGTFMFARLILISGPCDLPASASQASAGITGVSHARLIENFC 300

Qy 301 LFEMESHVTVQAGVQWPNLGLQPLPGKLRFSCLSLPSSWDYGHLPHPANFCIFIRGG 360
Db 301 LFEMESHVTVQAGVQWPNLGLQPLPGKLRFSCLSLPSSWDYGHLPHPANFCIFIRGG 360

Qy 361 VSPYLSGWSQTPDLR 375
Db 361 VSPYLSGWSQTPDLR 375

RESULT 3
US-09-964-667-2
; Sequence 2, Application US/09964667
; Publication No. US20030033621A1
; GENERAL INFORMATION:
; APPLICANT: de la Monte, Suzanne
; Wands, Jack R.
; TITLE OF INVENTION: Transgenic Animals and Cell Lines for
; Screening Drugs Effective for the Treatment or Prevention
; of Alzheimer's Disease
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox, P.L.L.C.
; STREET: 1100 New York Ave., Suite 600
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/964,667
; FILING DATE: 28-Sep-2001
; CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
```

QY	1	MEPSLLPLRECNCAISAHRNRLPGSSDSPASASPVAGITGMC	60	QY	181	VAGTKDMHYTWLIFIFINFLRQSLNSVTQAGVQWRNLGSLQLP	240
		THARLLLYFFFLVEMEF				PPGFKLFCSPSLSS	

Db 181 VAGTKDMHYTWLIFIFNFRLRQSLNSVTQAGVQWRNLGSLQPLPGFKLFCPSLLSS 240
Qy 241 WDYRRPPLANFFVFLVEMGTFMFARLLIISGPCDLPASASQSGAGITGVSHHARLIENFC 300
Db 241 WDYRRPPLANFFVFLVEMGTFMFARLLIISGPCDLPASASQSGAGITGVSHHARLIENFC 300
Qy 301 LFEMESHVSVTQAGVQWRNLGSLQPLPGFKLFCPSLLSSWDYGHLPHPANFCIFIRGG 360
Db 301 LFEMESHVSVTQAGVQWRNLGSLQPLPGFKLFCPSLLSSWDYGHLPHPANFCIFIRGG 360
Qy 361 VSPYLSGWSQTPDLR 375
Db 361 VSPYLSGWSQTPDLR 375

RESULT 6

US-10-146-130-2
; Sequence 2, Application US/10146130
; Publication No. US20030004107A1
; GENERAL INFORMATION:
; APPLICANT: AVERBACK, PAUL
; TITLE OF INVENTION: METHOD OF PREVENTING CELL DEATH USING SEGMENTS OF
; FILE REFERENCE: 59003.000007
; CURRENT APPLICATION NUMBER: US/10/146,130
; CURRENT FILING DATE: 2002-08-06
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 375
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-146-130-2

Query Match 100.0%; Score 2034; DB 14; Length 375;
Best Local Similarity 100.0%; Pred. No. 1e-184;
Matches 375; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MEFSLLLPRLECNAGISAHRNLRPLPGSSDSPASASPVAGITGCTHARLILYFFLVEMEF 60
Db 1 MEFSLLLPRLECNAGISAHRNLRPLPGSSDSPASASPVAGITGCTHARLILYFFLVEMEF 60
Qy 61 LHVGQAGLELPTSDPVSASQARYRTGHHARLCLANFCGRNRVSLMCPSPSPKQST 120
Db 61 LHVGQAGLELPTSDPVSASQARYRTGHHARLCLANFCGRNRVSLMCPSPSPKQST 120
Qy 121 CLSLPKCDYRRAAVPGLFIFLFLRHRCTTLTQDEVQWCDHSSLPQSTPEIKHPPASQ 180
Db 121 CLSLPKCDYRRAAVPGLFIFLFLRHRCTTLTQDEVQWCDHSSLPQSTPEIKHPPASQ 180
Qy 181 VAGTKDMHYTWLIFIFNFRLRQSLNSVTQAGVQWRNLGSLQPLPGFKLFCPSLLSS 240
Db 181 VAGTKDMHYTWLIFIFNFRLRQSLNSVTQAGVQWRNLGSLQPLPGFKLFCPSLLSS 240
Qy 241 WDYRRPPLANFFVFLVEMGTFMFARLLIISGPCDLPASASQSGAGITGVSHHARLIENFC 300
Db 241 WDYRRPPLANFFVFLVEMGTFMFARLLIISGPCDLPASASQSGAGITGVSHHARLIENFC 300
Qy 301 LFEMESHVSVTQAGVQWRNLGSLQPLPGFKLFCPSLLSSWDYGHLPHPANFCIFIRGG 360
Db 301 LFEMESHVSVTQAGVQWRNLGSLQPLPGFKLFCPSLLSSWDYGHLPHPANFCIFIRGG 360
Qy 361 VSPYLSGWSQTPDLR 375
Db 361 VSPYLSGWSQTPDLR 375

RESULT 7

US-10-092-934-10
; Sequence 10, Application US/10092934
; Publication No. US20030054990A1
; GENERAL INFORMATION:

; APPLICANT: AVERBACK, PAUL
; TITLE OF INVENTION: METHODS OF USING NEURAL THREAD PROTEINS TO TREAT TUMORS
; TITLE OF INVENTION: AND CONDITIONS REQUIRING THE REMOVAL OR DESTRUCTION OF
; FILE REFERENCE: 018792-0199
; CURRENT APPLICATION NUMBER: US/10/092,934
; CURRENT FILING DATE: 2002-06-05
; PRIOR APPLICATION NUMBER: 60/273,957
; PRIOR FILING DATE: 2001-03-08
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 375
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-092-934-10

Query Match 100.0%; Score 2034; DB 14; Length 375;
Best Local Similarity 100.0%; Pred. No. 1e-184;
Matches 375; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MEFSLLLPRLECNAGISAHRNLRPLPGSSDSPASASPVAGITGCTHARLILYFFLVEMEF 60
Db 1 MEFSLLLPRLECNAGISAHRNLRPLPGSSDSPASASPVAGITGCTHARLILYFFLVEMEF 60
Qy 61 LHVGQAGLELPTSDPVSASQARYRTGHHARLCLANFCGRNRVSLMCPSPSPKQST 120
Db 61 LHVGQAGLELPTSDPVSASQARYRTGHHARLCLANFCGRNRVSLMCPSPSPKQST 120
Qy 121 CLSLPKCDYRRAAVPGLFIFLFLRHRCTTLTQDEVQWCDHSSLPQSTPEIKHPPASQ 180
Db 121 CLSLPKCDYRRAAVPGLFIFLFLRHRCTTLTQDEVQWCDHSSLPQSTPEIKHPPASQ 180
Qy 181 VAGTKDMHYTWLIFIFNFRLRQSLNSVTQAGVQWRNLGSLQPLPGFKLFCPSLLSS 240
Db 181 VAGTKDMHYTWLIFIFNFRLRQSLNSVTQAGVQWRNLGSLQPLPGFKLFCPSLLSS 240
Qy 241 WDYRRPPLANFFVFLVEMGTFMFARLLIISGPCDLPASASQSGAGITGVSHHARLIENFC 300
Db 241 WDYRRPPLANFFVFLVEMGTFMFARLLIISGPCDLPASASQSGAGITGVSHHARLIENFC 300
Qy 301 LFEMESHVSVTQAGVQWRNLGSLQPLPGFKLFCPSLLSSWDYGHLPHPANFCIFIRGG 360
Db 301 LFEMESHVSVTQAGVQWRNLGSLQPLPGFKLFCPSLLSSWDYGHLPHPANFCIFIRGG 360
Qy 361 VSPYLSGWSQTPDLR 375
Db 361 VSPYLSGWSQTPDLR 375

RESULT 8

US-10-153-334-1
; Sequence 1, Application US/10153334
; Publication No. US20030096350A1
; GENERAL INFORMATION:
; APPLICANT: AVERBACK, PAUL
; TITLE OF INVENTION: PEPTIDES EFFECTIVE IN THE TREATMENT OF TUMORS AND OTHER
; TITLE OF INVENTION: CONDITIONS REQUIRING THE REMOVAL OR DESTRUCTION OF
; FILE REFERENCE: 59003-000006
; CURRENT APPLICATION NUMBER: US/10/153,334
; CURRENT FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: 60/293,156
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 375
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-153-334-1
Query Match 100.0%; Score 2034; DB 14; Length 375;

Best Local Similarity 100.0%; Pred. No. 1e-184;
Matches 375; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MEFSLLPLRLECNAGISAHRNRLPGSSDSPASASPVAGITGMCTHARLILYFPLVEMEF 60
Db 1 MEFSLLPLRLECNAGISAHRNRLPGSSDSPASASPVAGITGMCTHARLILYFPLVEMEF 60

Qy 61 LHVGQAGLELPTSDPVSASQSARYRTGHARLCLANFCGRNRVSLMCPSPWSPELKOST 120
Db 61 LHVGQAGLELPTSDPVSASQSARYRTGHARLCLANFCGRNRVSLMCPSPWSPELKOST 120

Qy 121 CLSLPKCDWYRRAAVPGLFILFHLRHCPTLTQDEVQWCHSSLOPSTPEIKHPPASASQ 180
Db 121 CLSLPKCDWYRRAAVPGLFILFHLRHCPTLTQDEVQWCHSSLOPSTPEIKHPPASASQ 180

Qy 181 VAGTKDMHYTWLIFIFNFRQSLNSVTQAGVQWRNLGSLQPLPGFKLFSCPSLLSS 240
Db 181 VAGTKDMHYTWLIFIFNFRQSLNSVTQAGVQWRNLGSLQPLPGFKLFSCPSLLSS 240

Qy 241 WDYRRPRLANFFVFLVEMGFTMFARLILISGPCDLPASASQSAGITGVSHHARLIENFC 300
Db 241 WDYRRPRLANFFVFLVEMGFTMFARLILISGPCDLPASASQSAGITGVSHHARLIENFC 300

Qy 301 LFEMESHVTOAGVQWPNLGLQPLPGPKLRFSCLSLPSSWDYGHLPHPANFCIFIRGG 360
Db 301 LFEMESHVTOAGVQWPNLGLQPLPGPKLRFSCLSLPSSWDYGHLPHPANFCIFIRGG 360

Qy 361 VSPYLSGWSQTPDLR 375
Db 361 VSPYLSGWSQTPDLR 375

RESULT 9
US-10-198-069-1
; Sequence 1, Application US/10198069
; Publication No. US20030096756A1
; GENERAL INFORMATION:
; APPLICANT: AVERBACK, PAUL
; TITLE OF INVENTION: PEPTIDES EFFECTIVE IN THE TREATMENT OF TUMORS AND OTHER
; TITLE OF INVENTION: CONDITIONS REQUIRING THE REMOVAL OR DESTRUCTION OF
; TITLE OF INVENTION: CELLS
; FILE REFERENCE: 59003.000009
; CURRENT FILING DATE: 2002-07-19
; PRIOR APPLICATION NUMBER: 60/306,161
; PRIOR FILING DATE: 2001-07-19
; PRIOR APPLICATION NUMBER: 60/306,150
; PRIOR FILING DATE: 2001-07-19
; PRIOR APPLICATION NUMBER: 60/331,477
; PRIOR FILING DATE: 2001-11-16
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: Patent in ver. 2.1
; SEQ ID NO 1
; LENGTH: 375
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-198-069-1

Query Match 100.0%; Score 2034; DB 14; Length 375;
Best Local Similarity 100.0%; Pred. No. 1e-184;
Matches 375; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MEFSLLPLRLECNAGISAHRNRLPGSSDSPASASPVAGITGMCTHARLILYFPLVEMEF 60
Db 1 MEFSLLPLRLECNAGISAHRNRLPGSSDSPASASPVAGITGMCTHARLILYFPLVEMEF 60

Qy 61 LHVGQAGLELPTSDPVSASQSARYRTGHARLCLANFCGRNRVSLMCPSPWSPELKOST 120
Db 61 LHVGQAGLELPTSDPVSASQSARYRTGHARLCLANFCGRNRVSLMCPSPWSPELKOST 120

Qy 121 CLSLPKCDWYRRAAVPGLFILFHLRHCPTLTQDEVQWCHSSLOPSTPEIKHPPASASQ 180
Db 121 CLSLPKCDWYRRAAVPGLFILFHLRHCPTLTQDEVQWCHSSLOPSTPEIKHPPASASQ 180

Qy 181 VAGTKDMHYTWLIFIFNFRQSLNSVTQAGVQWRNLGSLQPLPGFKLFSCPSLLSS 240
Db 181 VAGTKDMHYTWLIFIFNFRQSLNSVTQAGVQWRNLGSLQPLPGFKLFSCPSLLSS 240

Qy 241 WDYRRPRLANFFVFLVEMGFTMFARLILISGPCDLPASASQSAGITGVSHHARLIENFC 300
Db 241 WDYRRPRLANFFVFLVEMGFTMFARLILISGPCDLPASASQSAGITGVSHHARLIENFC 300

Qy 301 LFEMESHVTOAGVQWPNLGLQPLPGPKLRFSCLSLPSSWDYGHLPHPANFCIFIRGG 360
Db 301 LFEMESHVTOAGVQWPNLGLQPLPGPKLRFSCLSLPSSWDYGHLPHPANFCIFIRGG 360

Qy 361 VSPYLSGWSQTPDLR 375
Db 361 VSPYLSGWSQTPDLR 375

Best Local Similarity 100.0%; Pred. No. 1e-184;
Matches 375; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MEFSLLPLRLECNAGISAHRNRLPGSSDSPASASPVAGITGMCTHARLILYFPLVEMEF 60
Db 1 MEFSLLPLRLECNAGISAHRNRLPGSSDSPASASPVAGITGMCTHARLILYFPLVEMEF 60

Qy 61 LHVGQAGLELPTSDPVSASQSARYRTGHARLCLANFCGRNRVSLMCPSPWSPELKOST 120
Db 61 LHVGQAGLELPTSDPVSASQSARYRTGHARLCLANFCGRNRVSLMCPSPWSPELKOST 120

Qy 121 CLSLPKCDWYRRAAVPGLFILFHLRHCPTLTQDEVQWCHSSLOPSTPEIKHPPASASQ 180
Db 121 CLSLPKCDWYRRAAVPGLFILFHLRHCPTLTQDEVQWCHSSLOPSTPEIKHPPASASQ 180

Qy 181 VAGTKDMHYTWLIFIFNFRQSLNSVTQAGVQWRNLGSLQPLPGFKLFSCPSLLSS 240
Db 181 VAGTKDMHYTWLIFIFNFRQSLNSVTQAGVQWRNLGSLQPLPGFKLFSCPSLLSS 240

Qy 241 WDYRRPRLANFFVFLVEMGFTMFARLILISGPCDLPASASQSAGITGVSHHARLIENFC 300
Db 241 WDYRRPRLANFFVFLVEMGFTMFARLILISGPCDLPASASQSAGITGVSHHARLIENFC 300

Qy 301 LFEMESHVTOAGVQWPNLGLQPLPGPKLRFSCLSLPSSWDYGHLPHPANFCIFIRGG 360
Db 301 LFEMESHVTOAGVQWPNLGLQPLPGPKLRFSCLSLPSSWDYGHLPHPANFCIFIRGG 360

Qy 361 VSPYLSGWSQTPDLR 375
Db 361 VSPYLSGWSQTPDLR 375

RESULT 10
US-10-157-031-299
; Sequence 299, Application US/10157031
; Publication No. US20030108890A1
; GENERAL INFORMATION:
; APPLICANT: Baranova, A. V.
; APPLICANT: Yankovsky, N. K.
; APPLICANT: Kozlov, A. P.
; APPLICANT: Lobachev, A. V.
; APPLICANT: Krukovskaya, L. L.
; TITLE OF INVENTION: In silico screening for phenotype-associated expressed sequences
; FILE REFERENCE: 2760-103
; CURRENT APPLICATION NUMBER: US/10/157,031
; CURRENT FILING DATE: 2002-05-30
; NUMBER OF SEQ ID NOS: 415
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 299
; LENGTH: 375
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-157-031-299

Query Match 100.0%; Score 2034; DB 14; Length 375;
Best Local Similarity 100.0%; Pred. No. 1e-184;
Matches 375; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MEFSLLPLRLECNAGISAHRNRLPGSSDSPASASPVAGITGMCTHARLILYFPLVEMEF 60
Db 1 MEFSLLPLRLECNAGISAHRNRLPGSSDSPASASPVAGITGMCTHARLILYFPLVEMEF 60

Qy 61 LHVGQAGLELPTSDPVSASQSARYRTGHARLCLANFCGRNRVSLMCPSPWSPELKOST 120
Db 61 LHVGQAGLELPTSDPVSASQSARYRTGHARLCLANFCGRNRVSLMCPSPWSPELKOST 120

Qy 121 CLSLPKCDWYRRAAVPGLFILFHLRHCPTLTQDEVQWCHSSLOPSTPEIKHPPASASQ 180
Db 121 CLSLPKCDWYRRAAVPGLFILFHLRHCPTLTQDEVQWCHSSLOPSTPEIKHPPASASQ 180

Qy 181 VAGTKDMHYTWLIFIFNFRQSLNSVTQAGVQWRNLGSLQPLPGFKLFSCPSLLSS 240
Db 181 VAGTKDMHYTWLIFIFNFRQSLNSVTQAGVQWRNLGSLQPLPGFKLFSCPSLLSS 240

Qy 241 WDYRRPRLANFFVFLVEMGFTMFARLILISGPCDLPASASQSAGITGVSHHARLIENFC 300
Db 241 WDYRRPRLANFFVFLVEMGFTMFARLILISGPCDLPASASQSAGITGVSHHARLIENFC 300

Qy 301 LFEMESHVTOAGVQWPNLGLQPLPGPKLRFSCLSLPSSWDYGHLPHPANFCIFIRGG 360
Db 301 LFEMESHVTOAGVQWPNLGLQPLPGPKLRFSCLSLPSSWDYGHLPHPANFCIFIRGG 360

Qy 361 VSPYLSGWSQTPDLR 375
Db 361 VSPYLSGWSQTPDLR 375

```
RESULT 11
US-10-198-070-1
; Sequence 1, Application US/10198070
; Publication No. US20030109437A1
; GENERAL INFORMATION:
; APPLICANT: AVERBACK, PAUL
; APPLICANT: GEMMELL, JACK
; TITLE OF INVENTION: PEPTIDES EFFECTIVE IN THE TREATMENT OF TUMORS AND OTHER
; TITLE OF INVENTION: CONDITIONS REQUIRING THE REMOVAL OR DESTRUCTION OF
; TITLE OF INVENTION: CELLS
; FILE REFERENCE: 59003.000008
; CURRENT APPLICATION NUMBER: US/10/198,070
; CURRENT FILING DATE: 2002-07-19
; PRIOR APPLICATION NUMBER: 60/306,161
; PRIOR FILING DATE: 2001-07-19
; PRIOR APPLICATION NUMBER: 60/306,150
; PRIOR FILING DATE: 2001-07-19
; PRIOR APPLICATION NUMBER: 60/331,477
; PRIOR FILING DATE: 2001-11-16
; NUMBER OF SEQ ID NOS: 125
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 375
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-198-070-1

Query Match      100.0%; Score 2034; DB 14; Length 375;
Best Local Similarity 100.0%; Pred. No. 1e-184;
Matches 375; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MEFSLLLPRLECNCAISAHNRNLRPGSSDSPASASPVAGITGCTHARLILYFFLVEMEF 60
Db 1 MEFSLLLPRLECNCAISAHNRNLRPGSSDSPASASPVAGITGCTHARLILYFFLVEMEF 60

Qy 61 LHVGQAGLELPTSDPVSASQASARYRTGHARLCLANFCGRNRVSLMCPSPWSPKQST 120
Db 61 LHVGQAGLELPTSDPVSASQASARYRTGHARLCLANFCGRNRVSLMCPSPWSPKQST 120

Qy 121 CLSLPKCWDYRRAAVPGLFILFRLHRCPTLTQDEVQWCDHSSLOPSTPEIKHPPASQ 180
Db 121 CLSLPKCWDYRRAAVPGLFILFRLHRCPTLTQDEVQWCDHSSLOPSTPEIKHPPASQ 180

Qy 181 VAGTKDMHYTWLIFIFINFLRQSLNSVTQAGVQWRNLSLOLPPGFKLFSCPSLSS 240
Db 181 VAGTKDMHYTWLIFIFINFLRQSLNSVTQAGVQWRNLSLOLPPGFKLFSCPSLSS 240

Qy 241 WDYRRPPLANFFVFLVEMGFTMFARLILISGPCDLPASASQASAGITGVSHHARLIFNFC 300
Db 241 WDYRRPPLANFFVFLVEMGFTMFARLILISGPCDLPASASQASAGITGVSHHARLIFNFC 300

Qy 301 LFEMESHVTOAGVQWPNLGSLOLPPGLKRFSCLSLPSSWDYGHLPHPANFCIFIRGG 360
Db 301 LFEMESHVTOAGVQWPNLGSLOLPPGLKRFSCLSLPSSWDYGHLPHPANFCIFIRGG 360

Qy 361 VSPYLSGWSQTPDLR 375
Db 361 VSPYLSGWSQTPDLR 375

RESULT 12
US-10-755-889-410
; Sequence 410, Application US/10755889
; Publication No. US20040171823A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES ASSOCIATED WITH THE NF-kB
; TITLE OF INVENTION: PATHWAY
; FILE REFERENCE: D0284 NP
; CURRENT APPLICATION NUMBER: US/10/755,889
; CURRENT FILING DATE: 2004-01-13
; PRIOR APPLICATION NUMBER: U.S. 60/440,068
; PRIOR FILING DATE: 2003-01-14
```

```
; PRIOR APPLICATION NUMBER: U.S. 60/469,757
; PRIOR FILING DATE: 2003-05-12
; NUMBER OF SEQ ID NOS: 823
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 410
; LENGTH: 375
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-755-889-410

Query Match      100.0%; Score 2034; DB 16; Length 375;
Best Local Similarity 100.0%; Pred. No. 1e-184;
Matches 375; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MEFSLLLPRLECNCAISAHNRNLRPGSSDSPASASPVAGITGCTHARLILYFFLVEMEF 60
Db 1 MEFSLLLPRLECNCAISAHNRNLRPGSSDSPASASPVAGITGCTHARLILYFFLVEMEF 60

Qy 61 LHVGQAGLELPTSDPVSASQASARYRTGHARLCLANFCGRNRVSLMCPSPWSPKQST 120
Db 61 LHVGQAGLELPTSDPVSASQASARYRTGHARLCLANFCGRNRVSLMCPSPWSPKQST 120

Qy 121 CLSLPKCWDYRRAAVPGLFILFRLHRCPTLTQDEVQWCDHSSLOPSTPEIKHPPASQ 180
Db 121 CLSLPKCWDYRRAAVPGLFILFRLHRCPTLTQDEVQWCDHSSLOPSTPEIKHPPASQ 180

Qy 181 VAGTKDMHYTWLIFIFINFLRQSLNSVTQAGVQWRNLSLOLPPGFKLFSCPSLSS 240
Db 181 VAGTKDMHYTWLIFIFINFLRQSLNSVTQAGVQWRNLSLOLPPGFKLFSCPSLSS 240

Qy 241 WDYRRPPLANFFVFLVEMGFTMFARLILISGPCDLPASASQASAGITGVSHHARLIFNFC 300
Db 241 WDYRRPPLANFFVFLVEMGFTMFARLILISGPCDLPASASQASAGITGVSHHARLIFNFC 300

Qy 301 LFEMESHVTOAGVQWPNLGSLOLPPGLKRFSCLSLPSSWDYGHLPHPANFCIFIRGG 360
Db 301 LFEMESHVTOAGVQWPNLGSLOLPPGLKRFSCLSLPSSWDYGHLPHPANFCIFIRGG 360

Qy 361 VSPYLSGWSQTPDLR 375
Db 361 VSPYLSGWSQTPDLR 375

RESULT 13
US-10-910-173-2
; Sequence 2, Application US/10910173
; Publication No. US2005009041A1
; GENERAL INFORMATION:
; APPLICANT: Wands, Jack R
; APPLICANT: de la Monte, Suzanne M
; TITLE OF INVENTION: Inhibition of Neurodegeneration
; FILE REFERENCE: 21486-047
; CURRENT APPLICATION NUMBER: US/10/910,173
; CURRENT FILING DATE: 2004-08-02
; PRIOR APPLICATION NUMBER: US/09/872,968
; PRIOR FILING DATE: 2001-06-01
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 375
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-910-173-2

Query Match      100.0%; Score 2034; DB 17; Length 375;
Best Local Similarity 100.0%; Pred. No. 1e-184;
Matches 375; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MEFSLLLPRLECNCAISAHNRNLRPGSSDSPASASPVAGITGCTHARLILYFFLVEMEF 60
Db 1 MEFSLLLPRLECNCAISAHNRNLRPGSSDSPASASPVAGITGCTHARLILYFFLVEMEF 60

Qy 61 LHVGQAGLELPTSDPVSASQASARYRTGHARLCLANFCGRNRVSLMCPSPWSPKQST 120
Db 61 LHVGQAGLELPTSDPVSASQASARYRTGHARLCLANFCGRNRVSLMCPSPWSPKQST 120
```


Search completed: September 8, 2005, 11:17:14
Job time : 390 secs
